View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Columbia University Academic Commons

The Effect of Education on Earnings and Employment in the Informal Sector in South Africa

Izumi Yamasaki

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2012

© 2012
Izumi Yamasaki
All rights reserved

ABSTRACT

The Effect of Education on Earnings and Employment

in the Informal Sector in South Africa

Izumi Yamasaki

This dissertation analyzes the effects of schooling on earnings and employment in the informal private sector, compared with those in the public and formal private sectors, in South Africa. The estimations are conducted by race and gender as well to examine the difference between various subgroups. The research also examines the heterogeneity in returns to schooling.

The returns to schooling are estimated using two-stage least squares with multinomial two-step selection corrections to control for both endogeneity of schooling and sector sample selection bias. Quantile regression and piecewise linear spline function methods were applied to deal with the heterogeneity of the returns to schooling. Moreover, the Blinder-Oaxaca decomposition method was conducted to explore the contribution of returns to schooling and educational endowments to wage differentials between sectors. To analyze the effects of schooling on the probability of employment in different sectors of the economy, multinomial logit and probit models were estimated. The instrumental variable probit model was also used to control for the endogeneity of schooling.

Even after controlling for both endogeneity of schooling and sample selection bias, a significant difference was found in returns to schooling between the formal and informal private sectors. Returns to schooling in the formal private sector are higher than those in the informal private sector in general. These findings apply across race and gender, except for Whites and

Asians whose sample size in the informal sector was too small. They were robust even after controlling for district, industry, and occupation, and using different definitions of the informal sector. Therefore, the dissertation concludes that formal private and informal private labor markets are segmented in terms of returns to schooling in South Africa.

The study also showed a positive effect of schooling on employment in the public and formal private sectors versus employment in the informal private sector. The effect of schooling on employment in the formal private sector – when compared to that in the informal sector – was much stronger for females than males, except for Whites and Asians. The findings were robust after controlling for district and using different definitions of the informal sector.

TABLE OF CONTENTS

Chapter

LIST OF TABLES

Table

Table

Table

Table

Table

Table

Table

LIST OF FIGURES

Figure

ACKNOWLEDGEMENTS

I owe my deepest gratitude to my sponsor, Professor Francisco Rivera-Batiz for his unfailing commitment and guidance throughout my dissertation writing. It has been an honor to be his advisee and learn from his tremendous knowledge and insights. This study could not have been completed without his help. I also express my warmest gratitude to my advisors in the dissertation committee, Professor Mun C. Tsang, Professor Judith Scott-Clayton, Professor Douglas Ready, and Professor Ira Gang for their generosity and thought-provoking comments.

I am grateful to Makiko Shirota Nakamuro, who greatly contributed to my dissertation through shared research project. I truly appreciate her generous support. I would like to thank my peers in the Program in Economics and Education for their friendship and academic assistance. Especially, I am indebted to Mina Dadgar, Katia Herrera and Yuya Hoshino for their intellectual and moral support throughout my doctoral study. I would like to thank Susana Martinez Restrepo for sharing her experiences in dissertation preparation. I am thankful to many other students in my program for their feedback on my dissertation in the workshop. I am also grateful to participants in the workshops at the Center for the Study of Social Stratification and Inequality, Tohoku University and at Graduate School of Economics and the School of International Public Policy, Osaka University for their comments on this study.

Finally, I would like to thank my husband, my parents, my grandmothers, my sister, relatives and friends for their continuous support and encouragement.

DEDICATIONS

To my husband, Takashi Chikamoto, for his ceaseless love and care, intellectual assistance, and consistent patience.

To my parents, Satoru and Machiko Yamasaki, for their affection and continuous support of my education and career.

1. INTRODUCTION

This dissertation analyzes the effect of education on earnings and the probability of working in the informal sector, compared with the formal sector, in South Africa. The informal sector is usually associated with a specific form of the private sector where small and medium enterprises operate without legal recognition in developing countries.[1] Sub-Saharan Africa, where South Africa is located, has the largest informal sector in the world, with some estimates suggesting that as much a 75 percent of non-agricultural sector employment is accounted for by the informal sector.

Because such a significant portion of the population forms part of this sector, it is essential to understand the characteristics of its employment and how it differs from the formal sector. At the policy level, the government also needs to have information about the labor market outcomes in this sector, in order to plan adequate policies, whether in relation to poverty and income, or in terms of seeking ways to incorporate the informal sector into the formal. This research intends to provide a contribution to the understanding of the informal sector in South Africa.

1.1. Motivation for and Background of the Dissertation

In developing countries, a large part of the labor force is employed informally and a substantial fraction of GDP is generated from the informal sector. Not many comparative worldwide statistics are available on the size of informal employment and the economy because of the nature of the informal sector and employment. However, Charmes (2000) reported

[1]Defining the informal sector has been the subject of much debate. The definition of the informal sector and employment will be discussed in Chapter 2 of this study.

employment in the informal sector as a percentage of non-agricultural employment[2] by world regions in 1980-89 and in 1990-99, while Schneider (2006) calculated the size of the informal economy as a percentage of GDP[3] by group of countries in 1999/2000, 2000/2001, and 2001/2002 (see Tables 1.1 and 1.2). Both informal employment and the informal economy have a substantial share in developing countries, with an increasing trend. Between the 1980s and 1990s, the percentage of informal employment in non-agricultural employment increased from 39% to 43% in North Africa, from 68% to 75% in Sub-Saharan Africa, from 52% to 57% in Latin America, and from 53% to 63% in Asia (Charmes, 2000). Between 1999/2000 and 2001/2002, the size of the informal economy as a percentage of GDP slightly increased from 41% to 43% in Africa, 41% to 43% in Central and South America, 29% to 30% in Asia, 38% to 40% in transition countries, 32% to 33% in the South Pacific Islands, and 20% to 22% in Communist countries, while remaining around 16% in OECD countries (Schneider, 2006). Although the informal sector accounts for a large part of total employment, its relative contribution to GDP seems to be much smaller.

According to the traditional dual and segmented labor market view, the informal sector is seen as a disadvantaged, low-paying, and unprotected sector, as the last resort of employment for workers (Harris & Todaro, 1970; Lewis, 1954). Workers migrate from rural to urban areas in response to differences in expected earnings between the urban modern sector and the traditional agricultural sector (Harris & Todaro, 1970), and many of these workers enter the informal sector

---

[2]These statistics are in accordance with the definition of the informal sector from the 15th International Conference of Labour Statisticians in 1993 (International Labor Office [ILO], 1993).

[3]Schneider (2006) actually calls the informal economy a “shadow economy” instead of an informal economy. In his definition, the shadow economy includes all market-based legal production of goods and services that are deliberately concealed from public authorities.

because they fail to find jobs in the formal sector. The formal sector is thought to be overly regulated, leading the high unemployment rates (De Soto, 1989; Dickens & Lang, 1985;[4] Fields, 1975). Moreover, workers in the informal sector are thought to earn less than identical workers in the formal sector.

Some researchers, however, have challenged this traditional view, suggesting that a substantial share of informal activities may be the workers' voluntary choice, given their preferences, skills, and competing earnings prospects. This case has been made especially in reference to the self-employed in the informal sector of Latin America (Gindling, 1991; Jütting, Parlevliet, & Xenogiani, 2008; Maloney, 2004; Packard, 2007; Perry, Maloney, Arias, Fajnzylber, Mason, & Saavedra-Chanduvi, 2007). If labor markets are competitive, and not segmented, then wage convergence may be achieved in equilibrium for equivalent occupations in the two sectors and wage gaps would be explained by compensating differentials between the two (Bargain & Kwenda, 2009; Tannuri-Pianto, Pianto, & Arias, 2004).

Fields (1990) suggests that the informal sector could feature both types of workers: workers who have no choice but to work in the informal sector, and workers who choose to work in the sector. This third view, which combines both competitive and necessity views of informal work, has emerged only recently. Fields (1990, 2005) and others[5] claim that urban informal labor markets in developing countries consist of an "upper tier" and a "lower tier." The "upper tier" is the competitive part, with workers who voluntarily choose the informal employment, while the "lower tier" consists of individuals who cannot afford to be unemployed but also

---

[4]Dickens and Lang (1985) suggest the presence of labor market segmentation and dual labor markets in the U.S.

[5]For example, see Funkhouser (1997b) and Maloney (2004).

cannot find formal jobs.

De Soto (1989, 2000) and De Soto and Litan (2001) indicate that one of the reasons for the huge informal sector in developing countries is because being and staying formal is too costly in terms of time and money. There are too many regulations and procedures to start and run a formal business, and bureaucrats often ask them for bribes in return for providing permissions.[6] De Soto (1989) even points out that the possible reason why governments in developing countries have so many regulations is because government officials want to collect bribes.

1.2. Policies on the Informal Sector

Among policymakers, perspectives on the effect that the informal sector has on the economy and society have also been changing. Until the late 1980s, policymakers had usually ignored the informal sector, as Charmes (1990) points out. They thought the sector would diminish over time as countries develop. Many studies also argue that the informal sector is harmful. Loayza (1996) claims that the relative size of the informal economy reduces economic growth because of the low usage of public services by business firms and the increased number of activities using public services inefficiently. Also De Soto (2000, 2001) and De Soto and Litan (2001) suggest that the lack of legal property rights of people in the informal sector could be harmful on economic development. If people have no legal property rights, they cannot turn their assets into capital that can be used as collateral to start or invest in their businesses. A large informal sector has also been correlated with lower government revenues (Jütting & Laiglesia, 2009; United Nations Development Programme [UNDP], 2004) and higher taxes for firms in the

[6]Shleifer and Vishny (1993) say that such corruption is more costly than taxation.

formal sector (UNDP, 2004).

Despite views on the negative effects of the informal sector, governments in developing countries and international organizations began to support the informal sector in the late 1980s (Charmes, 1990; World Bank, 1987) because the sector continued to expand rather than diminish, and became an important source of employment in countries with high unemployment in the formal sector. Today, governments and international agencies recognize the informal sector as key for job creation, economic growth, and poverty reduction (Chen, 2007; Swedish International Development Cooperation Agency [SIDA], 2004). However, it should be noted that the ultimate goal for many governments is to expand the formal sector and diminish the role of the informal sector. Therefore, supporting the informal sector and its workers and encouraging them to become formal[7] have become essential goals for the economic development of developing countries.

With this background, this dissertation analyzes how education contributes to the earnings (productivity) of workers and the likelihoods of working in the informal sector in South Africa —compared to the formal sector— as a case study. In fact, education has been regarded as one of the most important components of labor market earnings (see the survey by Katz and Autor, 1999). It has also been identified as a key factor in economic growth. In their examination of the augmented Solow[8] model using actual cross-country data,[9] Mankiw, Romer, and Weil

---

[7]That is, workers in the formal sector with legal registration.

[8]The Solow model explains economic growth as a function of physical capital and amount of labor (Solow, 1956). Mankiw et al. (1992) expanded this model by adding human capital (education) as another explanatory variable of economic growth.

[9]Mankiw et al. (1992) used secondary school enrollments from the UNESCO yearbook as human capital in their production function. Many scholars are skeptical about the use of only secondary school enrollment as human

(1992) suggest that savings, education, and population growth are the major variables which account for cross-country variations in income per capita. More recent work has documented the importance of education for growth and development (Krueger and Lindahl, 2001, Goldin and Katz, 2007, and Rivera-Batiz, 2008).

1.3. Research Questions

The main research questions of this dissertation are:

1) To what extent does education affect the earnings of workers in the informal private sector? How does the effect of education on earnings differ between the formal and informal private sector, between the public and informal private sector, and between the public and formal private sector?
2) To what extent does education affect the probability of public, formal, and informal private sector employment?

For the first question, the focus is on calculating rates of return to education in the informal sector and to what extent they differ from those in other sectors. In developing countries, a large part of the labor force is employed in the informal sector, yet there is little information about the role education plays in determining wages. With these countries devoting much money to the expansion of their education system, especially for basic education, does such schooling pay-off in terms of higher wages in the informal sector? Or are earnings in the informal sector largely unrelated to schooling? This relates to the debate mentioned earlier on whether the informal sector is a sector of last resort for employment or not. The second question is also related to this issue, as whether the probability of employment in the informal sector responds to

---

capital. However, given the limited space in this study, a detailed discussion on how to measure human capital is not included here.

education or not is critical in determining what sector of the labor force it caters to.

1.4. Significance of the Study

The research in this dissertation can help policymakers 1) better understand the role played by education in determining labor market outcomes in the informal sector; 2) determine whether the sector is a supply of employment of last resort for highly unskilled and uneducated workers or, instead, a dynamic sector where highly-educated workers are employed in addition to low-educated workers; and 3) determine whether education can be a tool to mobilize workers from the informal sector into the formal sector, in South Africa and other developing countries. These policy issues are important for poverty reduction, job creation, and economic growth (Chen, 2007; SIDA, 2004). Finally, it is hoped that the research in this dissertation will contribute to the academic field of economics of education because few quantitative studies on the topic are available for South Africa. There are also fewer studies on this specific topic that involve rigorous identification strategies and econometric methods.

1.5 Structure of the Dissertation

This dissertation consists of 10 chapters. The first chapter explained the motivation for and background of the dissertation, policies on the informal sector, the research questions guiding the dissertation, and the significance of the study. The second chapter clarifies the definition of the informal sector and employment. The third chapter examines previous empirical studies on the impact of education on earnings of workers in the informal sector compared to the formal sector and on the effect of education on the probabilities of employment in both informal and formal sectors. The fourth chapter presents the theories on which the research is based. The fifth chapter describes the background of South Africa and the data used in the dissertation. The

sixth chapter presents the identification strategies utilized to estimate the returns to education in the informal and formal sectors and identification strategies used to estimate the effect of education on employment in different sectors. Then the descriptive statistics are explained in Chapter 7. Chapter 9 presents the estimation results of the return to education in the public, formal private, and informal private sectors. The policy recommendations, qualifications, and issues for future studies drawn from the analysis are also discussed in this chapter. Chapter 10 presents the estimation results of the effect of education on employments in different. The final chapter presents general conclusions, qualifications, and policy recommendations.

Table 1.1 *Informal Sector as % Non-agricultural Employment*

Region	1980-89	1990-99
North Africa	38.8	43.3
Sub-Saharan Africa	68.1	74.8
Latin America	52.3	56.9
Asia	53.0	63.0

Source: Charmes (2000), quoted in Beneria (2001), Table 6.

Table 1.2 *Average Size of the Informal (Shadow) Economy as % of Official GDP for Developing, Transition, and OECD Countries*

	1999/2000	2000/2001	2001/2002
Africa	41.3 (37)	42.3 (37)	43.2 (37)
Central and South America	41.1 (21)	43.1 (21)	43.4 (21)
Asia	28.5 (28)	29.5 (28)	30.4 (28)
Transition Countries	38.1 (25)	39.1 (25)	40.1 (25)
OECD Countries	16.8 (21)	16.7 (21)	16.3 (21)
South Pacific Islands	31.7 (10)	32.6 (10)	33.4 (10)
Communist Countries	19.8 (3)	21.1 (3)	22.3 (3)
Average Over 145 Countries	33.6	34.5	35.2

Source: Schneider (2006)
Note: The numbers in the parentheses shows the number of countries covered.

2. DEFINITIONS OF INFORMAL SECTOR AND INFORMAL EMPLOYMENT

This chapter discusses definitions of informal sector and informal employment. Given the illegality and ambiguity of the informal sector and employment, no consistent and universal definition of the informal sector and employment is available in the literature. The concept of the informal sector was first posed by the Kenya Report, written by the International Labor Office (ILO, 1972).[1] The ILO (1972) defines the informal sector as follows: (i) ease of entry, (ii) reliance on indigenous resources, (iii) family ownership of enterprises, (iv) small scale of operation, (v) labor-intensive technology, (vi) skills acquired outside the formal school system, and (vii) unregulated yet competitive market (p. 6).

In practice, the definition of the informal sector and employment differs from study to study. Table 2.1 presents the definition of informal workers (or sector) that will be used in the empirical research in this study. This definition often depends on data availability, and many papers use an incomplete definition of the informal sector because of a poor dataset. As Cole and Fayissa (1991) point out, definitions of informal sector/employment in early studies were often based on the size of the enterprises, the occupations of the workers, and the type of technology. Some studies defined the informal sector solely or heavily based on the size of the firm, i.e., the number of employees (Banerjee, 1983; Gong & van Soest, 2002; Pisani & Pagán, 2004; Tannuri-Pianto et al., 2004). For example, Pisani and Pagán (2004) defined informal workers as those who work in firms with five employees or less. Some studies used both size of firms and types of jobs to define informal workers (Funkhouser, 1996, 1997a, 1997b). A few studies in the

[1]Hart (1971), however, was the first to use the term.

literature review in this study (Packard, 2007; Saavedra & Chong, 1999) regarded informal workers as wage earners without a contract. Recent studies often define informal workers as those who work in firms (or are self-employed) without formal accounts, i.e., unregistered, not paying tax, and/or without social security (Kuepie, Nordman, & Roubaud, 2009; Marcouiller, Ruiz de Casilla, & Woodruff, 1997; Pagán & Tijerina-Guajardo, 2000; Tegoum, 2009; see Table 2.1 for more examples).

It is also important to note that informal workers usually include both self-employed and wage workers. The self-employed often accounts for the majority of workers in the informal sector. In all developing countries, self-employment comprises a greater share of informal employment than wage employment.[2]

There are sometimes subtle differences in terminology. Some countries make a distinction between "workers in the informal sector" and "informal workers". For example, Statistics South Africa (2010) defines informal workers as: workers in the informal sector plus employees in the formal sector and people working in private households who lack entitlements to basic benefits such as pension or medical aid and do not have a written contract of employment.[3] There is also the issue that some people may work in both the formal and informal sectors (Jütting et al., 2008). Finally, it is important to note that the informal sector and informal employment do not necessarily include criminal activities (SIDA, 2004).

This dissertation uses the so-called legalistic definition of the informal sector and informal workers. This is partly due to the fact that the research is based on data provided by a

[2]Self-employment accounts for 70% of informal employment in Sub-Saharan Africa, 62% in North Africa, 60% in Latin America, and 59% in Asia (SIDA 2004).

[3]Note, however, that no clear distinction between informal sector and informal employment will be used in the later chapters in this dissertation; rather, a clearer distinction will be used.

South African survey that asks respondents to identify whether they work in the informal sector or not, based on the legalistic definition. . If a respondent in this survey is working in the informal private sector, s/he is defined to be (1) working for a private company/institution whose employer (institution, business or private individual) is not registered for the Value Added Tax (VAT) to perform the activity, or (2) owns a business that is again not registered for the VAT. Most of domestic workers are included in the workers in the informal sector. Domestic workers are those who are paid to do domestic work, such as housekeepers, nannies, cooks, and gardeners. Their business is usually not registered. If their business is registered, they are included in the formal sector. If a respondent is working in the formal private sector, s/he is working for a private company/institution whose employer (institution, business or private individual) is registered (i.e. registered for VAT) to perform the activity, or owns a business which is again registered.

Table 2.1 *Definition of Informal Workers in Empirical Papers in the Literature Review*

Author(s) and Year Published	Country	Definition of the Informal Workers
Banerjee (1983)	Delhi, India	All workers working in private enterprises with less than 20 employees.
Gindling (1991)	San Jose, Costa Rica	Workers (i) whose occupations are not associated with government sponsorship;* (ii) who do not belong to a cooperative, union or professional organization; or (iii) who do not have post-graduate education.
Griffin & Edwards (1993)	Brazil	Workers without a signed workbook.
Funkhouser (1996)	Guatemala, El Salvador, Honduras, Nicaragua, Costa Rica.	(i) Workers in firms of four or fewer employees who are not professional, technical or administrative workers, and (ii) all self-employed workers.
Funkhouser (1997a)	Guatemala	Those (i) who are working in firms with four or fewer employees; and (ii) who are not owners, professional, technical or managerial workers.
Funkhouser (1997b)	El Salvador	(i) All self-employed, domestic, and family workers who are not employed in professional or technical occupations and any wage earners working in a firm with four or fewer employees, and (ii) owners who are excluded from both formal and informal sectors.
Marcouiller, Ruiz de Casilla, & Woodruff (1997)	Mexico, El Salvador, Peru	Workers who do not receive social security coverage or do not pay social security taxes as a condition of employment.
Saavedra & Chong (1999)	Peru	Informal wage earners are those without (i) a signed contract; (ii) union membership; (iii) health insurance or pension entitlement; (iv) vacation entitlement; and (v) a job in the public sector. The informal self-employed are those without (i) making any tax payment; and (ii) the registration with the tax authority.

Table 2.1 (continued)

Author(s) and Year Published	Country	Definition of the Informal Workers
Pagán & Tijerina-Guajardo (2000)	Mexico	Workers employed in a non-registered firm.
Tansel (2001)	Turkey	Wage earners who are not covered by any social security program and are self-employed (excluding professionals and technicians).
Gong & van Soest (2002)	Mexico	Workers (i) who work in firms with five or less employees or (ii) who work for their own account or manage a firm without employees.
Pisani & Pagán (2004)	Nicaragua	Workers who work in firms with five or fewer employees.
Tannuri-Pianto, Pianto & Arias (2004)	Bolivia	Informal salaried workers who work in establishments with 1-4 employees plus domestic employees.
Lassibille & Tan (2005)	Rwanda	Informal workers who work in the traditional sector.
Ewoudou & Vencatachellum (2006)	Cameroon	Workers who or whose employers are self-employed or not registered for tax purposes.
Packard (2007)	Chile	Non-contract wage earners and self-employed.
Stifel, Rakotomanana, & Celada (2007)	Madagascar	Workers without pension or social security.
Arias & Khamis (2008)	Argentina	Workers in a dependent employee relationship without social security contributions.
Kuepie, Nordman, & Roubaud (2009)	Seven urban cities in West Africa	Workers who work with production units with no fiscal or statistical identity or with no formal accountancy.
Tegoum (2009)	Cameroon	Informal sector workers who exercise their main job in a production unit and do not have a taxpayer number and/or keep formal accounts.

*These occupations are medical (doctors, nurses), social science, accounting and mathematics, law, the airline industry, communications, mail, mining, baking, insurance, electricity, gas, and sanitation.

3. LITERATURE REVIEW

3.1. Literature Review on Returns to Education in the Informal Sector Compared with the Formal Sector

This section summarizes previous empirical studies examining the returns to education in the informal sector compared with the formal sector. It then discusses empirical issues that have been unresolved in the previous studies. Most empirical studies use the Mincerian equation, i.e., rate of return to education approach, in order to estimate returns to education.

3.1.1. Previous empirical studies

This section summarizes and examines the identification strategies and findings from the existing literature. Most relevant previous studies were conducted in countries in Latin America or Africa.

3.1.1.1. Studies without correction for endogeneity

Early studies did not correct for endogeneity of either sector or education (Banerjee, 1983; Funkhouser, 1996, 1997b; Griffin & Edwards, 1993; Smith & Metzger, 1998). Most of this literature found a positive return to education in the informal sector, but one that was usually much smaller than that in the formal sector. Only Banerjjee (1983) found no significant difference in estimated returns to education between the informal and formal sectors. The results of these early studies, however, may be biased because they did not control for selectivity effects or endogeneity.

3.1.1.2. Studies with correction for endogeneity of sector allocation

Most previous studies on returns to education in the formal and informal sectors, and the determinants of earnings in these two sectors, have controlled for endogeneity of sector

allocation or sample selection bias in the sectoral allocation of workers. Many studies have used the Heckman or Lee type two-step model to correct for sectoral allocation bias.

The previous empirical studies which did correct for sector selection bias also found substantial returns to education in the informal sector. Some studies found that the return to education was greater in the formal sector than in the informal sector (Argentina: Arias, & Khamis, 2008; Guatemala: Funkhouser, 1997a; Nicaragua: Pisani & Pagán, 2004; Peru: Saavedra & Chong, 1999). Others, however, suggested greater returns to education in the informal sector than in the formal sector (Bolivia: Tannuri-Pianto et al., 2004; Cameroon: Ewoudou & Vencatachellum, 2006; Madagascar: Stifel, Rakotomanana, & Celada, 2007; Rwanda: Lassibille & Tan, 2005; Turkey: Tansel, 2001), especially for secondary and vocational education (Ewoudou & Vencatachellum, 2006; Stifel et al., 2007; Tannuri-Pianto et al., 2004). In fact, Tansel (2001) and Lassibille and Tan (2005) showed greater returns to education in the informal sector than in the formal sector, except for tertiary education. .

As mentioned earlier, many studies used the Heckman or Lee type two-step selection model to correct for sectoral allocation bias. Funkhouser (1997a) and Arias and Khamis (2008) employed the Heckman correction to control for selection bias of sectoral allocation in the formal and informal sectors. Gindling (1991), Saavedra and Chong (1999), Lassibille and Tan (2005), and Ewoudou and Vencatachellum (2006) utilized the Lee two-step selection model to correct for selection bias for more than two sectors, including the informal sector. Stifel et al. (2007) applied a similar selection correction using the multinomial logit model proposed by Bourguignon, Fournier, and Gurgand (2007). Pisani and Pagán (2004) conducted the switching regression model to estimate earnings in multiple sectors.

Lassibille and Tan (2005) and Tannuri-Pianto et al. (2004) also used the wage

decomposition method. They decomposed the earning gaps into differences in endowments of skills and differences in return to skills.

3.1.1.3. Studies with corrections for endogeneity of both sector allocation and education

Many studies on the returns to education, especially in developed countries, have corrected for the endogeneity of education. The most popular methodologies include using family background controls (e.g., Griliches, 1979), siblings and twins models (e.g., Ashenfelter & Krueger, 1994), and instrumental variables (e.g., Angrist & Krueger, 1991). In addition, as noted above, a large number of studies have corrected for sectoral allocation between the formal and informal sectors. However, few studies have corrected for endogeneity of both sectoral allocation and education

The paper written by Kuepie et al. (2009) is probably the most rigorous study in this field because the authors attempted to correct for endogeneity of both sector allocation and education in the context of estimating returns to education in the formal and informal sectors while also taking into account the heterogeneity of returns to education. They examined returns to education in the public sector, the formal private sector, and the informal private sector using comparable labor force surveys in urban cities in West African countries. For the sector selection bias, the authors employed the multinomial logit model proposed by Lee (1983) to estimate the probability of employment in the formal versus informal sectors, the first equation in a two-stage method.[1] To deal with the endogeneity of the education variable, they first conducted the control

[1]Kuepie et al. (2009) note that researchers such as Vijverberg (1993), Dahl (2002), and Bourguignon et al. (2007) have criticized Lee's method because it is based on a strong assumption about the point distribution of error terms of the equations of interest. However, Kuepie et al. (2009) found that the existing alternative methods did not

function (CF) method,[2] using father's schooling and main occupation as instruments. They also used family background information differently by introducing it directly into the earnings functions. Finally, the authors estimated both a piecewise linear spline function and an ordinary marginal returns to education function. A piecewise linear spline function enables the relationship between education and earnings to vary across different parts of the education distribution.

Kuepie et al. (2009) found that the return to education rises once an endogenous education variable is corrected for. This effect was particularly true in the informal sector. In general, education made significant impacts on individual earnings in the informal private sectors of these West African cities, even at high levels of schooling. In most of the West African cities in their sample, the public sector had greater returns to education than the formal private sector, followed by the informal private sector. The public sector had the greatest marginal return to education, while the informal private sector had the smallest return to education in most cities. However, the return to schooling in the informal private sector was still substantial. The results showed convex returns to education in all cities and sectors, including the informal sector. While the return to primary schooling in the informal private sector tended to be weaker than in the formal private sector, the return to lycée (upper secondary education) in the informal private sector was greater than in the public sector in most cities and in the formal private sector in some cities.

Gong and van Soest (2002) used a dynamic random effects panel data model consisting

---

appear any more efficient given the small size of their sectoral sub-samples; therefore, they chose Lee's correction method in which the correction terms are easier to interpret.

[2]See Garen (1984), Wooldridge (2002), and Söderbom, Teal., Wambugu, and Kahyarara (2006), as mentioned by Kuepie et al. (2009). Chapter 7 in this study explains the method in greater detail.

of separate wage equations for the two sectors to estimate wage equations in both the formal and informal sectors in urban Mexico. They found that returns to education were positive in both sectors, but much higher in the formal sector than in the informal sector for both men and women. They found large positive wage differentials between the formal and informal sectors for the higher education levels, and small or even negative differentials for the medium and low education levels.

Tegoum (2009) also addressed the selection biases of both sector and education. He estimated rates of return to primary education completion and the first cycle of secondary education completion within the informal sector in Cameroon. The author used the two matching methods (propensity score matching and Epanechinikov kernel matching) to correct for the endogeneity of education and the selection correction method according to Bourguignon et al. (2004) to correct for sectoral allocation. The author found a significant positive impact of schooling on the income of informal sector workers. However, we cannot tell if the returns to education in the informal sector are greater or smaller than those in the formal sector because the author did not estimate returns to education in the formal sector.

3.1.1.4. Studies dealing with the heterogeneity in returns to education

To deal with heterogeneity of returns to education, Arias and Khamis (2008) employed the marginal treatment effects according to the recently developed models by Heckman and Vytlacil (2001, 2005) and Heckman, Urzua, and Vytlacil (2006). Moreover, Kuepie et al. (2009) used a piecewise linear spline function in addition to controlling for the endogeneity of sector and education. Tannuri-Pianto et al. (2004) estimated quantile earnings equations, as originally proposed by Koenker and Bassett (1978), to analyze the heterogeneous effect of education on earnings in different sectors as well as correcting for the endogeneity of sector allocation using

Lee's two-step selection method: the approach similar to that proposed by Fitzenberger (2003).

3.1.2. Summary and Methodological Issues Unresolved

The previous empirical studies on returns to education found substantial returns to education in the informal sector. More than half of the studies in this review found that returns to education were greater in the formal sector than in the informal sector. However, a number of studies also found greater returns to education in the informal sector than in the formal sector for all or some levels of education, especially for secondary and vocational education. The mixed results on the relative magnitude of returns to education in the formal and informal sectors could be due to different labor market characteristics in each country. However, these estimates could be still biased because most of them did not control for the endogeneity of education. The results could also be biased because of some measurement errors. For example, measurement errors might be found in the sector variable because defining the informal employment and sector is complicated. Attenuation bias might also be present because of measurement errors in the reported years of schooling. Measurement errors may appear in reported earnings, especially for non-salaried workers in the informal sector, as suggested by Kuepie et al. (2009).

Most of the reviewed studies used a log of "hourly" earnings, which is advisable (see Table 3.1). As Griliches (1977) suggests, wage rates per hour or week can be used as a dependent variable for return to schooling estimates when they are available. Griliches explains that annual earnings or income "confound market transactions with issues of labor-leisure choice and the more transitory effects of unemployment" (p. 3).

As for the use of education as an independent variable, some of these studies used a years of schooling continuous variable, while others used a set of dummy variables for level of education (see Table 3.2). Years of schooling explains how one more year of schooling affects

the wage rate. However, we have to assume constant returns to one additional year of schooling at any level. By contrast, including dummy variables for each level of education can estimate different rates of return to different levels of education, but not the marginal effect of one more year of schooling. Most of the empirical studies reviewed here tended to have only one of two different types of variables, but it would be a good idea to conduct two regressions using two different types of variables for schooling. Moreover, not many studies account for the heterogeneity of the return to education except for Kuepie et al. (2009), Arias and Khamis (2008), and Tannuri-Pianto et al. (2004). For future studies, it will be important to consider the heterogeneous returns to education by sector.

3.2. Literature Review on the Effect of Education on Informal and Formal Employment

The previous section reviewed the literature on the effects of education on workers' earnings; however, studying the education and earnings of workers in the informal sector is not enough. One must also examine the determinants of the probability of employment in that sector relative to the formal sector. As a matter of fact, this analysis is required in order to correct for possible sample selection bias in the earnings equation, to ensure that estimated rates of return to education are not influenced by the sample of workers who chooses to be employed in the informal sector relative to the formal sector.

To analyze the sectoral allocation or choice between sectors, one can use either a probit or a binary logit model if there are only two categories, while the multinomial logit model is used with more than two categories.

3.2.1. Previous empirical studies

As just noted, most studies examining the effects of education on the probability of informal employment are often carried out jointly with the analysis of rates of return to education

in the informal sector because these studies employ the Heckman or Lee two-step selection model, using either the probit or multinomial logit model to estimate the propensities of working in each sector as a first stage in correcting earnings estimates for sector allocation bias. However, no study has controlled for the endogeneity of education in the probability equation.

Most prior studies have found that education is positively associated with formal employment and negatively associated with informal employment in general (see Table 3.3). The more education a worker attains, the more likely he or she will work in the formal sector (Argentina: Arias & Khamis, 2008; Cameroon: Tegoum, 2009; 14 other studies).

As for methodologies applied, studies considering only two categories (informal employment and formal employment) used either the binary logit model (Banerjee, 1983) or probit model (Arias & Khamis, 2008; Funkhouser, 1996, 1997b; Marcouiller et al., 1997, Pagán & Tijerina-Guajardo, 2000; Tansel, 2001). Studies looking at more than two categories used the multinomial logit model.[3] More than half of the studies in this literature review analyzed the probabilities of working in more than two sectors, such as the public, formal private, and informal private sectors (Banerjee, 1983; Ewoudou & Vencatachellum, 2006; Gindling, 1991; Lassibille & Tan, 2005; Packard, 2007; Saavedra & Chong, 1999; Stifel et al., 2007; Tannuri-Pianto et al., 2004; Tegoum, 2009).

3.2.2. Summary and methodological issues unresolved

The existing empirical studies on the effects of education on the probability of formal and informal employment share similar findings. The more education a worker attains, the more likely he or she will work in the formal sector. But the fact that these studies have ignored the

[3]Theoretically, the multinomial probit model exists, but it is virtually impossible to use it because the computation used in the model is very complicated.

endogeneity of education raises questions about their results. It is possible that workers who had more education tended to have higher innate abilities or higher socio-economic status and were more likely to work in the formal sector because of this. Thus, future studies should include stronger identification strategies.

Table 3.1 *Summary of Literature on the Return to Education in the Formal and Informal Sectors*

Author(s) and Year Published	City & Country	Data	Group	Endogeneity of Sector Allocation	Endogeneity of Education	Results
Gindling (1991)	San Jose, Costa Rica	Survey conducted in San Jose	Public, formal private, & informal private	Used the Lee (1983) selection correction at first but ended up not using it.	N/A	Positive marginal RTE in the informal private sector, but it is the highest in the public sector, followed by the private formal sector.
Funkhouser (1997a)	Guatemala	HHS	Public, formal private, & informal private	Heckman (1979) selection correction	N/A	Positive marginal RTE in the informal sector. It is greatest in the formal private, followed by the informal private.
Saavedra & Chong (1999)	Peru	LSMS	Formal and informal wage earners, & formal and informal self-employed	Lee (1983) selection correction	N/A	Positive marginal RTEs for informal wage earners and self-employed. Marginal RTE is the highest for formal wage earners, followed by formal self-employed, and then informal self-employed.
Tansel (2001)	Turkey	HHES	Formal., informal & self- employed	Lee (1983) and Trost & Lee (1984) selection correction	N/A	Greater RTEs for the informal wage earners than formal ones, except for university education. Much lower RTE for self-employed than other groups.
Gong & van Soest (2002)	Mexico	UES	Formal & informal	Dynamic random effects panel data model		Positive RTEs in the informal sector, but much greater RTEs in the formal sector.
Pisani & Pagán (2004)	Nicaragua	LSMS	Formal & informal	Switching regression model	N/A	Positive marginal RTE in the informal sector but greater marginal RTE in the formal sector.

Table 3.1 (continued)

Author(s) and Year Published	City & Country	Data	Group	Endogeneity of Sector Allocation	Endogeneity of Education	Results
Tannuri-Pianto, Pianto, & Arias (2004)	Bolivia	LFS and HHSs	Formal., informal., & self-employed	Quantile regression	N/A	Greatest RTEs to secondary and technical education in the informal sector. Similar high returns to university education for all sectors.
Lassibille & Tan (2005)	Rwanda	HLCS	Public, formal private, & informal	Trost & Lee (1984) selection correction	N/A	Greatest RTEs in the informal private except for higher education.
Ewoudou & Vencatachellum (2006)	Cameroon	HHS	Public, formal private, & informal	Trost &d Lee (1984) selection correction	N/A	Positive RTEs in the informal private but the public sector has the greatest RTEs in general. Greater RTEs in the formal private compared to informal private except for upper secondary and vocational education.
Stifel, Rakotomanana, & Celada (2007)	Madagascar	HHS	Public, formal private, & informal private	Bourguignon et al.'s (2007) selection correction	N/A	Positive RTEs in the informal sector but the public sector has the greatest RTEs. The returns to lower and upper secondary education are greater in the informal private sector than the formal private sector.
Arias & Khamis (2008)	Argentina	HHS and other data	Formal & informal salaried workers, and the self-employed	Heckman (1979) selection correction	N/A	Much greater RTEs in the formal sector for both secondary and tertiary education.

Table 3.1 (continued)

Author(s) and Year Published	City & Country	Data	Group	Endogeneity of sector allocation	Endogeneity of education	Results
Kuepie, Nordman, & Roubaud (2009)	7 urban cities in West Africa	Urban HHSs	Public, formal private, & informal private sector	Lee (1983) selection correction	CF method & father's characteristics control variables	Positive RTEs in the informal sector. Greater RTEs in the public sector, followed by the formal private and then the informal private in most cities.
Tegoum (2009)	Cameroon	EESI	Public, private formal., non-agricultural informal., & agricultural informal	Bourguignon et al.'s (2004) selection correction	Propensity score matching & Epanechinikov kernel matching	Positive RTEs for informal sector workers.

Notes: EESI: Survey on Employment and the Informal Sector; HHES: Household Expenditure Survey; HHS: Household Survey, HLCS: Household and Living Conditions Survey; LSMS: Living Standards Measurement Survey; UES: Urban Employment Survey; RTE: return to education; Marginal RTE: return to an additional year of education

Table 3.2 *Dependent and Education Independent Variables of Wage Equation*

Author(s) and Year Published	Dependent Variable	Education Independent Variable(s)
Banerjee (1983)	Log monthly earnings	Level of education dummies
Gindling (1991)	Log hourly wage	Years of schooling
Griffin & Edwards (1993)	Log monthly earnings	Years of schooling Years of schooling dummies
Funkhouser (1996)	Log of weekly earnings	Years of schooling
Funkhouser (1997a)	Log of monthly income	Years of schooling
Funkhouser (1997b)	Log of monthly income	Years of schooling
Marcouiller et al. (1997)	Log of hourly wage	Years of schooling
Saavedra & Chong (1999)	Log of hourly income and benefits	Years of schooling
Pagán & Tijerina-Guajardo (2000)	Log of hourly wage	Years of schooling
Tansel (2001)	Log of hourly wage	Level of education dummies
Gong & van Soest (2002)	Log of hourly wage	Level of education dummies
Pisani & Pagán (2004)	Log of weekly earnings	Years of schooling (no wage equation)
Tannuri-Pianto, Pianto, & Arias (2004)	Log of hourly earnings	Level of education dummies
Lassibille & Tan (2005)	Log of hourly earnings	Years of schooling Level of education dummies
Ewoudou & Vencatachellum (2006)	Log of hourly earnings	Level of education dummies
Stifel, Rakotomanana, & Celada (2007)	Log of real daily earnings	Level of education dummies
Arias & Khamis (2008)	Log of labor income per hour in the main occupation	Level of education dummies (only secondary & tertiary education dummies)
Kuepie, Nordman, & Roubaud (2009)	Log of hourly earnings	Years of each level of education times each level of education dummy
Tegoum (2009)	Log of hourly income	First School Leaving Certificate (FSLC = primary), General Certificate of Education Ordinary Level (GCE-OL = lower secondary)

Table 3.3 *Summary of Literature on the Effect of Education on the Probability of Formal and Informal Employment*

Author(s) and Year Published	City & Country	Data	Group	Methodology	Results
Banerjee (1983)	Delhi, India	Collected by the author	Formal., informal., nonwage employment	Multinomial logit, binary logit	Matriculated education and above has positive effect on formal employment. Education is not a significant determinant on wage and nonwage employment into the informal sector. Education has a significant influence on mobility from the informal to the formal sector except for below middle education and graduate education and above, which negatively correlates with mobility from the informal sector into the formal sector.
Gindling (1991)	San Jose, Costa Rica	Survey conducted in San Jose	Public, formal private, & informal private	Multinomial logit	Positive for public employment and negative for informal employment, compared with private-formal employment.
Funkhouser (1996)	Guatemala, El Salvador, Honduras, Nicaragua, Costa Rica	HHSs	Formal & informal	Probit	Negative effect of education on informal employment.
Funkhouser (1997b)	El Salvador	National HHS	Formal & informal	Probit	Positive effect of education on the probability of formal employment, compared with informal employment across gender.
Marcouiller, Ruiz de Casilla, & Woodruff (1997)	Mexico, El Salvador, Peru	HHSs	Formal & informal	Probit	Positive effect of education on the probability of formal employment, compared with informal employment across gender and countries.

Table 3.3 (continued)

Author(s) and Year Published	City & Country	Data	Group	Methodology	Results
Saavedra & Chong (1999)	Peru	LSMS	Formal., informal wage and informal self-employment	Multinomial logit	More schooling decreases the probability of informal wage employment and self-employment, but increases the probability of formal wage and self-employment.
Pagán & Tijerina-Guajardo (2000)	Mexico	ENEU (1987, 1993)	Formal & informal	Probit	Education decreases informal employment, compared with formal employment.
Tansel (2001)	Turkey	HHES	Formal., informal, & self-employed	Probit	(Men) Any levels of education decrease the probability of covered and uncovered wage employment and self-employment, compared with other employment (including public sector). Greatest negative effect of education on the probability of uncovered wage employment. (Women) Any level of education reduces the probability of informal wage employment and increases the probability of formal wage employment and self-employment (except for university education).
Pisani & Pagán (2004)	Nicaragua	LSMS	Formal & informal	Switching regression model	Education across years is one primary determinant of formal sector participation, compared with the informal sector participation.
Tannuri-Pianto, Pianto, & Arias (2004)	Bolivia	LFS and HHSs	Formal., informal, & self-employed	Multinomial logit	Primary and higher levels of education (compared with the excluded category less than basic) decrease the probability of informality. Any levels of education (compared with the excluded category) decrease the probability of self-employment.
Lassibille & Tan (2005)	Rwanda	HLCS	Public, formal and informal wage, & informal wage, self-employed, unpaid family worker and out of the labor force	Multinomial logit	More education decreases the probability of working in the informal sector and increases the probability of working in the public and formal private sectors.

Table 3.3 (continued)

Author(s) and Year Published	City & Country	Data	Group	Methodology	Results
Ewoudou & Vencatachellum (2006)	Cameroon	HHS	Public, formal private, & informal	Multinomial logit	More education decreases the probability of working in the informal sector and increases the probability of working in the public and formal private sectors.
Packard (2007)	Chile	LSMS-type HHS	Formal., informal., & self-employed	Multinomial logit	Negative for informal employment. Any levels of education except for primary and incomplete secondary decrease the probability of working in the informal sector.
Stifel, Rakotomanana, & Celada (2007)	Madagascar	HHS 2005	Public, formal private & informal private	Multinomial logit	The effect of education on the probability of informal employment is mixed.
Arias & Khamis (2008)	Argentina	HHS and other data	Formal salaried workers, informal salaried workers, & self-employed	Probit	Formal vs. informal: positive effect of secondary and tertiary education on the probability of working as formal salaried workers. Formal vs. self-employed: no significant differences in the probability of working in both groups. Informal vs. self-employed: negative effect of secondary and tertiary education on the probability of working as informal salaried workers.
Tegoum (2009)	Cameroon	EESI	Public, private formal., non-agricultural informal., & agricultural informal	Multinomial logit (odds ratio)	Any level of education is associated with much greater probability of working, compared with no education in the public and private formal sectors, compared with the informal sector. But very small effect of education on the probability of working in the formal sector.

Notes: EESI: Survey on Employment and the Informal Sector; HHES: Household Expenditure Survey; HHS: Household Survey, HLCS: Household and Living Conditions Survey; LSMS: Living Standards Measurement Survey; UES: Urban Employment Survey.

4. THEORETICAL FRAMEWORK

This dissertation seeks to estimate whether returns to education differ between the formal and informal sectors. It also examines how education affects the probability of working in the informal and formal sectors. But what does economic theory tell us about these issues?

This section first discusses dual and competitive market theories and draws implications from them in terms of the returns to education in the informal and formal sectors. It then reviews the Roy model and human capital theory, upon which the dissertation is primarily based. Analyzing returns to education in different sectors is closely connected to the endogeneity of education and sector sample selection. Thus, it is important to review the Roy model, which models occupational choice, as well as human capital theory, which models educational choice based on rates of return to schooling.

4.1. Dual and Competitive Labor Market Theories and Their Implications for Returns to Education

As explained in the introduction, three theories have been posed to explain wage differentials between the informal and formal labor markets. The first is the dual labor market theory, which asserts that the informal and formal sectors are segmented so that the wages for identical workers in both sectors are different (Harris & Todaro, 1970; Lewis, 1954). The returns to education in the two sectors are different, with the formal sector having higher rates of return to education and the informal sector remaining as an employer of last resort where rates of return to education are low.

A second theory assumes that labor markets are competitive (Gindling, 1991; Jütting et al., 2008; Maloney, 2004; Packard, 2007; Perry et al., 2007). In this case, wages for identical

workers are the same in the two sectors, but observed wage differentials are due to different characteristics and preferences of workers in the two sectors. The third theory combines both the competitive and segmented views on the informal labor market (e.g., Fields, 2005). The informal labor market consists of the competitive part which has the same returns to education as the formal sector, or even higher returns to education than the formal sector, and the segmented part, which has lower returns to education, than the formal sector.

4.2. Roy Model

This dissertation employed the Roy model as the theoretical underpinning for analyzing the inter-sectoral allocation of workers. The implications of the theory were used in the empirical analysis to model the selection correction equation in the estimation of the returns to schooling and also in the choice equation in the analysis of the effects of schooling on employment in different sectors.

Roy (1951) examined the effects of self-selection in different occupations based on unobservable comparative advantages and preferences among individuals, and analyzed the effects on the distribution of earnings in different occupations.[1] The following model explains how individuals choose a sector based on the Roy model in a general case which could have more than two sectors. The equations are formulated by reference to Dahl (2002), Arias and Khamis (2008), and Rankin, Sandefur, and Teal (2010).[2]

Suppose there are several types of occupations: s. Workers choose their occupation by comparing the utility W_s they derive from each occupation, which is given by the sum of the

---

[1]Roy's (1951) paper does not include any equation, but his theory has been applied by many researchers.

[2]See Dahl (2002) and Rankin et al. (2010) for a more detailed discussion on the theoretical and mathematical set-up of the general Roy model.

income Y_s and non-pecuniary benefits in the sector e_s net of costs c_s (pecuniary and non-pecuniary) of sector participation. Adapting a latent index formulation, we have:

(4.1) $W_{si}^{*} = Y_{si} + e_{si} - c_{si} = Z_i^{'}\gamma_s + \eta_{si}$

where W_{si}^{*}, an individual's utility of a particular occupational choice, depends linearly on the vectors of observed Z (e.g., human capital) and unobserved characteristics η (e.g., preference for work and abilities) of the worker i. A worker chooses an occupation s when the net benefits of working in the sector are positive.

Y_{si} is observed if and only if category s is chosen. This happens when

(4.2) $W_{si} > \max_{j \neq s}(W_{ji})$

Given this sorting, we will observe earnings in sector s; by definition, no earnings are yielded the unemployed. If we define

(4.3) $\varepsilon_s = \max_{j \neq s}(W_{ji} - W_{si}) = \max_{j \neq s}\left[Y_{ji} + e_{ji} - c_{ji} - (Y_{si} + e_{si} - c_{si}) \right] = \max_{j \neq s}(Z\gamma_{ji} + \eta_{ji} - Z\lambda_{si} - \eta_{si})$,

then under this definition, condition (5.2) is equivalent to $\varepsilon_s < 0$

(4.4) $Y_s = X\beta_s + u_s$

where s is the sector which the individual chooses.

(4.5) $E(Y_s \mid X, s = 1) = X_i\beta + E(u \mid X_i, s_i = 1)$

(4.6) $E(Y_s \mid X, s = 1) = X_i\beta + E(u_i \mid X_i, \varepsilon_i < 0)$

In applying the model above, if there are two types of sectors s: 1 for formal employment and 2 for informal sectors, we can draw the two equations below:

(4.7) $Y_1 = X\beta_1 + u_1$

(4.8) $Y_2 = X\beta_2 + u_2$

where X is a subset of Z and (u_1, u_2) are freely correlated and independent of some components of Z, the 'instruments.' The u_s can depend on η_s in a general way. This model can be applied to the case with more than two sectors.

4.3. Human Capital Theory

To model the basic determinants of earnings, and possible biases arising from the role played by ability, this dissertation uses the traditional human capital model developed by Becker, Mincer and Card, among many others.. Traditional human capital theory asserts that human capital, i.e., skills obtained through education and training, raise the productivity of workers and increase wages (Becker, 1964; Mincer, 1958). Becker (1964) and Mincer (1974) provide theoretical and empirical explanations for the human capital theory. This section reviews the basic human capital model.

The model below is based on Card (1995, 1999), who developed a model of optimal schooling investments.[3] He assumed that individuals (or their parents) try to maximize utility, which is a function of income and schooling.

(4.9) $U(y, S) = \ln(y) - f(S)$

$= \ln[g(S)] - f(S)$

[3]See Card (1995, 1999) for a more detailed discussion on the theoretical and mathematical modeling on returns to schooling.

where $y = g(S)$ represents the observable relationship between earnings (y) and schooling (S), and $\ln[g(s)]$ and $f(S)$ are increasing convex functions representing the (log) benefits and costs of schooling. To maximize utility in equation (5.9), optimal schooling (S^*) must satisfy the first-order condition,

(4.10) $g'(S)/g(S) = f'(S)$

where marginal benefits $(g'(S)/g(S))$ are equal to marginal costs $(f'(S))$.

Implementing this model empirically requires choosing functional forms for the marginal benefits and costs of schooling. Here, individual heterogeneity is considered in the costs of (or tastes for) schooling, $h(S)$, and in the economic benefits of schooling $y'(S)/y(S)$.

The marginal benefit (MB_i) of schooling and the marginal cost (MC_i) are represented by

(4.11) $MB_i = g'(S_i)/g(S_i) = b_i - k_1 S$ and

(4.12) $(MC_i) = f'(S_i) = r_i + k_2 S_i$

where b_i and r_i are random variables with means $\bar{b}$ and $\bar{r}$, and k_1 and k_2 are non-negative constants. The optimal level of schooling is

(4.13) $S_i^* = \dfrac{b_i - r_i}{k}$

where $k = k_1 + k_2$.

Finally, Eq. (5.11) implies a model for log earnings of the form

(4.14) $\ln y_i = \alpha_i + b_i S_i - \dfrac{1}{2} k_1 S_i^2$

where α_i is a person-specific constant which integrates individual unique characteristics.

Assuming that b_i and r_i are symmetrically distributed, the population regression coefficient is

(4.15) $\rho = (1-\alpha)\bar{b} + \alpha\bar{r}$

where $\alpha = \frac{k_1}{k} - \lambda$ and $\lambda = \frac{\sigma_b^2 - \sigma_{br}}{(\sigma_b^2 - \sigma_{br}) + (\sigma_r^2 - \sigma_{br})}$

With the simplest model of endogenous schooling, where individuals seek to maximize the discounted present value of earnings at fixed individual-specific discount rates (i.e., $k_2 = 0$), $\alpha = 1 - \lambda$, conventionally estimated return to schooling is

(4.16) $\rho = \lambda\bar{b} + (1-\lambda)\bar{r}$

Using the fact that $\bar{\beta} = \bar{b} - k_1\bar{S}$, Eq. (5.15) can be written as

(4.17) $\rho = \bar{\beta} + \lambda(\bar{b} - \bar{r})$

Therefore, in an OLS regression of log earnings on schooling, the estimate of the average marginal return to schooling will be upward-biased. The bias will get bigger as σ_b^2 (the variance in ability) gets larger relative to σ_r^2 (the variance in discount rates). $\lambda(\bar{b} - \bar{r})$ represents an endogeneity bias created because people with higher marginal returns to education seek more schooling.

4.4. Theory of Allocation of Talent

Relevant to this dissertation is also the work of Murphy, Shleifer and Vishny (1991), who modeled how individuals choose their occupation, and suggested that the allocation of talent (based on education or ability) among various sectors of the economy has significant effects on growth.

In their paper, Murphy, Shleifer and Vishny discuss the allocation of talent between two

sectors. According to them, the most competent people are drawn into the sector (sector 1) with less diminishing returns (higher elasticity of output with respect to labor). The returns to education are higher in this sector. But if the sector lacks dynamic externalities (such as in fostering innovation and technical change), then the absorption of educated workers in the sector will slow down economic growth. This is the case when highly educated workers are employed in low-productivity growth sectors, which may include government employment or, sometimes, informal sector employment that lacks the links to the formal sector necessary for developing and spreading new products or technologies. There is therefore a potentially inefficient outcome in their theoretical model: the competent people become employed in sector 1, where positive externalities are none or few, while the persons employed in sector 2, where dynamic externalities exist, are much less competent. As a result, the economy grows at a slower rate. This model is relevant for discussions involving the informal sector.

5. CONTEXT OF SOUTH AFRICA AND DATA DESCRIPTION

This chapter describes the South African context and the datasets used in the dissertation.

5.1. Context of South Africa

This section reviews the history, economy, employment, income, education and other social aspects of South Africa. It especially analyzes the history of apartheid and its impacts on racial inequality in employment, income and education. To sum, South Africa is upper middle income country with high unemployment rate and great inequality in income, employment, and education to name a few. Even now, Whites are privileged in any aspects while Blacks tend to be much more disadvantaged. These are considered to be mostly the legacies of apartheid.

5.1.1. Demography and geography of South Africa

South Africa is a middle size country in terms of population and surface area, with a high rate of urban population. It is also a multiracial country. The total population of South Africa is 50 million (2010, see Table 5.1), which is approximately 6% of the Sub Sahara Africa region. From 2000 to 2010, population growth has decreased from 2.4% to 1.4% (Table 5.2). Population growth within this decade was lower than the Sub Saharan Africa average but greater than BRIC[1] countries and OECD countries. The urban population as a percentage of total population in South Africa was 61.7% in 2010, which is much higher than the average of Sub Saharan Africa and slightly higher than the average of upper middle income countries (see Table 5.1). The South African population consists of four basic types of races: Blacks (Africans), Whites, Coloreds and Indians/Asians. Blacks account for 79.5% of population, Whites 9.0%, Colored

[1] BRIC countries are Brazil, Russia, India, and China and South Africa. This group of countries is slightly different from BRICS countries which includes South Africa.

9.0% and Indians/Asians 2.5% in 2011[2]. The term, African is sometime used instead of Black. In order to be consistent, Black will be used throughout this study. To note, the term, Colored is an established terminology in South Africa and will be used to express the people who are half White and half Black.

South Africa is the southernmost country of the Sub Saharan Africa region. Its surface area is 1,219.1 thousand square kilometers, which accounts for 5% of the region and one tenth of the surface of the United States (see Table 5.1). South Africa is made of nine provinces: Western Cape, Northern Cape, Eastern Cape, KwaZulu-Natal, Free State, North West, Gauteng, Mpumalanga, and Limpopo. South Africa has three capitals: Pretoria, the executive capital where the executive branch of the government is; Cape Town, the legislative capital which is the home to the Parliament; and Bloemfontein, the judicial capital which houses the Supreme Court.

5.1.2. Economy and unemployment in South Africa

South Africa is an upper middle-income country[3] and the largest economy in the Sub Saharan Africa region albeit with a high poverty rate and economic inequality. In many ways, the South African economy has done well since apartheid ended in 1994. The country holds valuable natural resources and a relatively developed infrastructure, while efficient and modern financial institutions has supported economic growth (World Bank, 2012 April). Its local currency is the rand (sign: R). The official exchange rate (local currency per US dollar) was 7.32 rand in 2010. Its Gross Domestic Product (GDP) in 2010 was 187.2 billion, approximately 35%

---

[2] Statistics South Africa (2011).

[3] This category is according to World Bank (2009).

of GDP in the Sub Saharan Africa region. Its GDP per capita was $3,746 in 2010[4], which is higher than most of other Sub Saharan African countries and also higher than other BRIC countries except for Brazil (see Table 5.3).

Macroeconomic management has been creditable; the inflation rate continues to be in the targeted 3-6 percent range and budgetary and debt outcomes have been world class (World Bank, 2012 April). Before the abolishment of apartheid, the GDP growth rate in South Africa stagnated, but since 1994 it has been decent and stable, although it is considered modest by international comparisons. The growth rate ranged between a healthy 3 to 6 percent except for the years of 1997, 2009 and 2010 when the global financial crises affected its economy. Per capita GDP in South Africa increased by 50% between 1999 and 2008 in South Africa (World Bank, 2010). The GDP growth rate in South Africa in 2009 was -1.7%, due to the global financial crisis yet it had recovered to 2.8% by 2010. GDP growth in South Africa has been historically lower than the Sub Saharan Africa average and the majority of BRIC countries in the 2000s (see Table 5.4); however, growth spillovers to the rest of the region are amazingly large by international standards (World Bank, 2010). An additional percentage point of South African GDP growth is associated with 0.4 to 0.9 percent of GDP growth in the rest of Sub Saharan Africa, independent of regional shocks (World Bank, 2010).

In spite of a positive economic performance, the country has been suffering from a high unemployment rate and great inequality in income, employment, and education. Table 5.5 shows the unemployment rate in South Africa and other selected countries. Since 2000, the unemployment rate in South Africa has been higher than 20%; it has been the highest in Sub

---

[4] The number is in constant year 2000 US$.

Saharan African region and is one of the highest unemployment rates in the world. The most recent unemployment rate is 23.8% (2009), which suggests the issues of unemployment is chronic as it has not been resolved after many years. Furthermore, the unemployment rate is very high especially for Black people. This issue will be examined more in detail in a later section.

The informal sector also accounts for a substantial part of employment and economy, albeit it is relatively smaller than other African countries. South Africa's informal economy as a percentage of total GDP was estimated to be 29.5% in 2002/03 (Schneider, 2006), while its share of total employment was estimated to be 22.3% in 2002 (Davies & Thurlow, 2009). Other studies, using alternative definitions of the informal sector, have estimated the size of the informal economy and employment to be even larger.

5.1.3. History and policies of the apartheid

South Africa once held an infamous racial discrimination policy called apartheid (1948 to 1994). As mentioned above, South Africa still faces the negative legacies of the apartheid. It is important to review apartheid, its policies and its impact on South African society until now as this plays a prominent role on today's effect of schooling on earnings and likelihood of working in the informal sector.

Apartheid, meaning 'apartness" or "separateness" in the Afrikaan and Dutch languages, is a policy of separating people by race, with regard to places of residence, schools, workplaces, and even cemeteries (Clark & Worger, 2004). The policy was introduced in 1948 by the National Party government. However, racial discrimination did exist in South Africa before 1948. It had existed since the beginning of Dutch colonization of the country in 1652 in the form of slavery. Even after the abolishment of slavery in the 1830s, racial discrimination continued as European settlement expanded (Clark & Worger, 2004). South Africa was not very different from other

African and Asian countries which were colonized until the end of the Second World War. However, South Africa went backward after the war and instead implemented apartheid, while indigenous people won their independence from Europeans in other Asian and African countries (Clark & Worger, 2004).

Many scholars in the 1970s and 1980s argued that it was not history or culture, but economic concerns, that were the basis for apartheid. They suggested this extreme racial discrimination was motivated by white business owners who believed a massive low-paid African workforce would allow them to make exceptional profits (Clark & Worger, 2004). Other scholars have argued that white workers and farmers also benefited economically from racially discriminatory laws which protected them from competition with African workers and producers (Clark & Worger, 2004).

Under apartheid, Blacks were excluded from all rights normally associated with those of citizens. Numerous laws were passed under the racist government in the 1950s. For example, the 1953 Reservation of Separate Amenities Act imposed segregation on all public facilities, including post offices, beaches, stadiums, parks, toilets, cemeteries, busses and trains (South Africa History Online, accessed on May 4, 2012). By the 1980s, many Blacks in South Africa were no longer legal citizens of that country but rather categorized as foreigners (Clark & Worger, 2004).

During apartheid, a large share of African population were segmented to tribal homelands and urban townships. At the beginning of apartheid, Blacks was removed from towns and white-owned farms they used to live and moved to 10 bantustans (known as the Black African homelands) to form the ethnically homogeneous towns. Also, Colored voters were removed form common voters roll in 1950s and later lost indirect parliamentary representation.

Black voters lost indirect parliamentary representation as well, but later obtain franchise in bantustans (Seekings & Nattrass, 2005).

There were also racial discriminations in welfare and benefits. Although the government developed segregated townships in the 1950s to move Blacks from their original habitation areas, in 1960s the government prohibited further development (Seekings & Nattras, 2005). Although primary education for Black children expanded rapidly in the 1960's, the education they received was of poor quality and insufficiently funded. The 1953 Bantu Education Act enforced the control of Black education by linking tax receipts from blacks to public expenditure on their education: in 1975, the average expenditure on education for Whites was more than 15 times larger than that for Blacks (Thomas, 1996). In 1970s, the government increased social spending on Blacks and greatly expanded secondary education for Black children. However, according to Thomas (1996), Whites continued to enjoy better funded schooling. For example, education spending per student is R4 for Whites and R1 for Blacks. Asians and Colored received better schooling than Blacks although it was also worse than Whites; education spending per student was R3 for Asians and R2 for Coloreds. Also development was permitted in African townships in 1970s (Seekings & Nattras, 2005). Racial discrimination was also applied to labor market. In 1953, the Industrial Conciliation Act removed African workers the right to strike and limited their representations to "workers committees". A color bar was implemented to prevent Coloreds and Blacks from working in white-dominated workplaces particularly in manufacturing (Seekings & Nattras, 2005).

The apartheid era has been criticized by both domestic and international communities and it ended in 1994 when the political party called African National Congress (ANC) won a landslide victory. The party has won victories in all four subsequent democratic elections (World

Bank, 2012 April). The country had a relatively peaceful transition from racist doctrine to inclusive democracy, which is thought to be one of the greatest achievements in late twentieth century (Welsh, 2009; World Bank, 2012 April). However, the great racial inequalities due to apartheid still exist. The inequalities in income, employment and education are explained more in detail later.

5.1.4. Tax revenue, tax system and public expenditure in South Africa

This section looks at public revenue and expenditure. The total public expenditure in South Africa in 2010 was 2669.5 billion rand (in current local currency), equivalent to 202.6 billion US dollars (constant 2000 US$) and approximately 100% of the GDP[5]. South Africa has been successful in collecting tax from upper income groups compared to other countries, although it has a substantial informal economy. Table 5.6 shows that tax revenue as a percentage of GDP in South Africa is higher than the Sub Sahara African average and much higher than BRIC and selected OECD countries.

South Africa's income tax system is progressive. Let us look at the income tax rate in 2007 when the Labor Force Survey used by this study was conducted in Table 5.7. For example, if a person under 65 years old earns R112,500 (categorized as the lowest income group), the estimated income tax is R12,510, which is 11% of total income. If a person under 65 years old earn R350,000 (the third income group), the estimated income tax is 27% if total income. If a person under 65 years old earns R700,000 (the richest income group), the estimated income tax is 32% of total income. For the richest group, the tax rate get closer to 40% as the income

[5] The data is from World Bank WDI & GDF database (accessed on May 4, 2012).

increases. Considering the highest tax rate in the United States is 35%[6], the income tax system in South Africa is more progressive than the one of the United States.

Public spending on education in South Africa has been more than 5% of GDP which is higher than the average of Sub Saharan Africa, and higher than BRIC countries in the 2000s (see Table 5.8). Also, its public spending on education as a percentage of all government spending has been more than 17% since 2000; this is similar to the Sub Saharan African average and higher than BRIC and selected OECD countries (Table 5.9). In 2010, public spending as a percentage of GDP and as a percentage of government spending in 2010 was 6% and 19.2% respectively. These data show that South Africa also has been investing heavily in education.

5.1.5. Unions, wages and minimum wage law in South Africa

Next, trade unions, industrial council, and their impact on wages in South Africa are reviewed. It is important to explain them since the minimum wage law affects employment numbers in both the formal and informal sectors and trade unions have substantial power in determining minimum wage and wage takes. According to Butcher and Rouce (2001), unions were able to raise wages by 10% for Whites and 20% for Blacks in 1995. However, unions seemingly contribute to the inflexibility of labor market in South Africa. Union Black workers earn 60% more than nonunion Black workers (Schults & Mwabu, 1998). These union wage premiums could affect informal employment numbers.

In South Africa, the Department of Labour sets minimum wage laws accordingly by sector and occupation. The sectoral determinations are retail and whole sale trade, domestic workers, farm workers, forestry workers, taxi operators, private security workers, hospitality

---

[6] United States Internal Revenue Service (2011).

sector workers and contract cleaners (Development Policy Research Unit: DPRU, 2008). However, the real minimum wages are estimated to be lower than the minimum wage stipulated by the laws (DPRU, 2008). As mentioned, even the minimum wage for domestic workers who are categorized as workers in the informal sector exists. Dinkeleman and Ranchhod (2010) found that the minimum wage law for domestic workers substantially increased the wage of domestic workers but not their employment.

5.1.6. Education system and basic education statistics in South Africa[7]

This section overviews the basic information and statistics on the South African education system. South Africa has 6 years of primary education (the first cycle of general and compulsory education), 3 years of lower secondary education (the second cycle of general and compulsory education, 3 years of upper secondary education or upper secondary training, and tertiary education (about 4 years for Bachelors, 7 or more years for postgraduate degrees) (see Table 5.10).

Early childhood education is provided by provincial Departments of Education but its provision is still limited and preschool education, for children aged 3, is dominated by the private sector. The Department of Education began to focus on providing one year of early childhood education prior to primary schooling for five-year-olds. The government has decided to provide one-year compulsory early childhood education (the reception year of compulsory education, called grade R) and is planning to enroll all children aged 5 in grade R by 2014.

Primary education is from grade 1 to 6. Children start their primary education when they are 7 years old. Primary and lower secondary education is compulsory and called general

---

[7] The information in first four paragraphs in this section is based on UNESCO (2010, August).

education. Lower secondary education is from grade 7 to 9. Upper secondary education is from grade 10 to 12. It includes either general upper secondary education or technical (upper) secondary education. Technical secondary education is provided by technical centers, high schools and vocational schools. Vocational programs requires three years in general, leading to the National Certificate (Vocational), NCV3. NCV1 and NCV2 are awarded if students complete one or two years of vocational education.

Tertiary and higher education is provided by colleges, technikons, universities of technology, and universities. Colleges for educational professions, i.e. teaching, often have 3-year diploma programs, while nursing colleges offer 4-year diploma programs. Agricultural colleges have 1 to 2 year certificate programs and 3 year diploma program. Universities usually give students bachelors after 3-4 years of study (5 years for architecture and law; 5 and a half years for veterinary medicine; 6 years for medicine and surgery). Master degrees require additional 1 to 2 years of study, while doctoral degrees require a minimum of 2 years.

Tables 5.11 to 5.14 display the gross and net enrollment rates for primary and secondary education for South Africa, the Sub Saharan Africa Region and selected countries. Looking at the gross enrollment rate for primary education, South Africa achieved universal primary education more than two decades ago: it has been more than 100% since the late 1980s. However, net enrollment rates for primary education have been decreasing. The primary net enrollment rate once reached 95% in 1995, but it has continuously decreased since then: it is now smaller than most of other middle income countries in the region and BRIC countries. The primary net enrollment rate in 2009 was in fact 85%. For secondary education, both gross and net enrollment rates have been increasing and they are higher than many other Sub Saharan middle income

countries and BRIC countries. The secondary gross enrollment rate was 86% in 2000 and it was 94% as of 2009. The net enrollment rate for secondary was 62% in 2000, and 72% in 2009.

5.1.7. Existing segregations and inequalities

Even after the abolishment of apartheid, Blacks were still separated from Whites to some extent in terms of habitation area, employment and education. For example, Capetown is the metropolis where Black people are not the majority, and it is still deeply divided (Polgreen, 2012).

5.1.7.1. Poverty, and inequalities in unemployment and income

South Africa is considered to be one of the most unequal countries in the world. Table 5.15 shows the Gini coefficient in South Africa, selected neighboring African countries, BRIC countries, and selected OECD countries. The Gini coefficient in South Africa in 2009 was 63.1, the highest in the world. Surprisingly, inequality in income has been increasing since apartheid ended. Table 5.16 shows the poverty rate (poverty headcount ratio) at the national poverty line for South Africa and neighboring middle income countries and BRIC countries. The poverty headcount ratio in South Africa has decreased from 31% in 1995 to 23% in 2006. The poverty rate in South Africa during the mid and late 2000s was lower than most of selected neighboring countries and Brazil and India. If we look at the poverty rate at $2 per day (PPP)[8], the story changes a little bit (see Table 5.17). The poverty rate in South Africa has been decreasing but it is higher than those at the national poverty line. For example, the poverty rates in South Africa in 2006 and 2009 were 35.7% and 31.3% respectively. The poverty rate at $2 a day was much lower than neighboring middle income countries but higher than Brazil, China and India.

---

[8] PPP means purchasing power parity.

Table 5.18 displays the poverty rate by race. Both in 1996 and 2001, the poverty rate for Blacks was around 35% which is much higher than those of other races, while the one for Whites was very small (1%). The high poverty rate for Blacks is due to the high unemployment rate for Blacks.

The economy has been unable to generate sufficient jobs, especially for Blacks. Table 5.19 shows the unemployment rate by race. The unemployment rate for the Black population is the highest, and three times as high as that for the Indian population and 4.5 times as high as that for the White population. The unemployment rate is 29.5% for Black, 22.5% for Coloreds, 10.1% for Indian, and 6.4% for White. This trend is similar for young first-time job seekers. Employment rates for school-leaving first-time entrants into the labor market were 29% for Blacks, while it is 75% for Whites (Table 5.20). Looking at the average monthly wage by race in Table 5.21, Blacks again had the lowest average monthly wage, while the White population had the highest. The average monthly wage for Whites was more than 4 times as high as that for Blacks in 2008.

Moreover, the majority of workers in the informal sector in the country are Black. According to Valodia, Lebani, Skinner, and Devey (2006), Blacks account for 89.3% of people in the informal sector, while Whites account for 4.7%. In addition, women are over-represented in the informal labor market. Estimates are that as much as 55.9% of workers in the informal sector are women, according to the South Africa Labor Force Survey in September, 2007.

Blacks were discriminated against in employment under apartheid (Bhorat, Lundall, & Rospabé, 2002; Chamberlain & van der Berg, 2002) and the statistics for recent years show that the great racial inequalities still persists.

5.1.7.2. Inequality in education

Also, disparities between Whites and Blacks in education have continued to exist after apartheid was abolished (O'Gorman, 2007). It is also important to note again that the explicitly discriminatory policies of apartheid were applied to education prior to 1994. But the legacy of apartheid has also had a great impact on inequality in education between Whites and Blacks until the present (Oosthuizen & Bhorat, 2006).

If we look at the percentage of the population aged 20 and older by race and by level of education (Table 5.22), Blacks have the highest percentage of population with no schooling; they represent the lowest percentage of people with grade 7 completed and higher. In 2006, the percentage of people with no schooling was 13.1% for Blacks, 4.4% for Colored, 2.9% for Indian, and 0.1% for Whites. The percentage of people with some form of primary schooling in 2006 was 16.9% for Blacks, 15.9% for Coloreds, 5.1% for Indian, and 0.6% for Whites. The percentage of people who completed grade 7 and higher was 70% for Blacks, 79.7% for Coloreds, 92% for Indian and 99.2% for Whites.

Table 5.23 shows the gross enrollment rate in public higher education institutions according to race. Although, enrollment rates in primary and secondary education is now closer to those of Indians and Whites, and the enrollment rate in higher education has been rapidly improving for Blacks and Coloreds, the enrollment rate in higher education for them is still very low compared to Indians and Whites. The gross enrollment rate in public higher education institution was recorded at 12% for Blacks and Coloreds, 51% for Indian and 59% for Whites in 2006.

Inequality in the quality of schooling also continues to be an issue. For example, According to van der Berg (2002), who surveyed 16,000 schools in seven stages, 95.2% of

schools were occupied by students from a single population group. It is important to note that the average teacher-student ratio of schools with Black students is 35.1 while it is 25.6 in schools with White students. Also, the percentage of qualified teachers for Blacks and Coloreds were 93.9% and 92.2%; lower than those for Indians and Whites (98.1% and 99.5%), in 2008 (See Table 5.24). However, the percentage of qualified teachers has been dramatically improving last two decades for Blacks and Coloreds.

5.1.8. Other human development challenges

There are some other significant human development challenges in South Africa. First, gender discrimination has been persistent. For example, the female homicide rate is six times as high as the world average; 50% of women are killed by their partners (Lancet, Health in South Africa, 2009 cited in World Bank, 2010). Second, South Africa has serious health issues. South Africa has 0.7% of world's population, but 17% of the world's HIV/AIDS cases and 24% of the global burden of HIV-related TB (World Bank, 2010). Life expectancy at birth has continued to decrease, 59.9 years in 1995 to 52 years in 2010[9]. Third, South Africa has substantial urban/rural disparities (Leibbrand, Woolard, & Woolard). This is connected to racial disparities because the large share of the rural population is Black.

5.2. Data Description

This dissertation uses South African labor force surveys that were conducted twice a year between 2000 and 2007 (in February and September from 2000 to 2002, and in March and September from 2003 to 2007) by Statistics South Africa. According to Statistics South Africa

---

[9] World Bank WDI & GDF database (accessed on May 4, 2012).

Labor Force Survey Reports,[10] all nine provinces of South Africa were included. The data were collected by survey questionnaire. The number of those who were of working age in the sample population ranged from 67,000 to 74,000 for these surveys. The number of households covered in the surveys was over 30,000.

The survey was designed to be a nationally representative sample. For example, for the Labor Force Survey of September 2007, the survey randomly drew the sample of Primary Sample Units (PSUs) for the Master Sample. The Master Sample is a multi-stage stratified sample. Enumeration Areas (EAs) with a household count of less than 25 were omitted from the census frame that was used to get the sample of PSUs. Other omissions from the Master Sample frame are all institutional EAs except workers' hostels, convents, and monasteries. EAs in the census database that were found to have less than 60 dwelling units were pooled together. The sample size of PSUs was 3,000. The explicit strata were the 53 district councils (DCs). The 3,000 PSUs were allocated to these using the power allocation method. The PSUs were sampled using probability proportional to size principles. The measure of size used was the number of households in a PSU based on census 2001 (Statistics South Africa, 2007b).

The datasets have a sector variable which indicates whether an individual was working in formal or informal sectors. This variable was created based on questions asked of individuals about "What type of occupation and job task are" and "In which sector the organization/business/enterprise/branch where a respondent works is". For the latter question, sector means either formal sector or informal sector (including domestic work). Also, respondents are told that the "formal sector employment is where the employer (institution,

---

[10]Statistics South Africa, 2003, 2004a, 2004b, 2004c, 2005a, 2005b, 2006a, 2006b, 2007a, 2007d.

business or private individual) is registered to perform the activity," whereas the "informal sector employment is where the employer is not registered" (Statistics South Africa, 2007c, p. 24).

The survey also asked additional questions which could be used to create other variables on informal employment. For example, the September 2007 survey asked whether a respondent's employer contributed to any pension or retirement fund, and how many employees there were in an organization, business, enterprise or branch where a respondent worked. The survey also asked whether the respondent's organization, business, enterprise or branch was a registered organization, paid for medical aid, was registered for VAT, and paid income tax (Statistics South Africa, 2007b).

As we can expect, the dataset has some missing values of variables of interest. For example, 28.3% of workers do not have their earnings information. However, the author could not find a significant difference in the distributions of education variables between those who report earnings and those who did not report earnings. Therefore, it is still possible to conduct the return to education analysis for workers who report their earnings.

Table 5.1 *Population Indicators and Surface Area for South Africa, Sub Saharan Africa Region and Upper Middle Income Countries: 2007 and 2010*

	South Africa		Sub Saharan Africa		Upper Middle Income Countries	
	2007	2010	2007	2010	2007	2010
Population (millions)	48.3	50.0	793.3	854.1	2,402.1	2,452.1
Urban population (% of total population)	60.3	61.7	36.0	37.4	55.4	57.4
Population growth (%)	1.1	1.4	2.5	2.5	0.7	0.7
Surface area (thousand sq. kim)	1,219.1		24,270.7		59,327.8	

Source: World Bank World Development (WDI) & Global Development Finance (GDF) database (accessed on May 4, 2012).

Table 5.2 *Population Growth for South Africa, Sub Saharan Africa Region and Selected Countries: 1990-2010*

	1990	1995	2000	2005	2007	2008	2009	2010
South Africa	2.0	2.2	2.5	1.1	1.1	1.1	1.1	1.4
Sub Saharan Africa	2.8	2.7	2.6	2.5	2.5	2.5	2.5	2.5
(Neighboring middle income countries)								
Angola	2.8	3.0	3.0	3.3	3.0	2.9	2.8	2.8
Botwaqna	3.0	2.5	1.7	1.3	1.4	1.4	1.4	1.3
Ghana	2.7	2.6	2.4	2.4	2.4	2.4	2.4	2.4
Namibia	3.9	2.9	2.4	1.8	1.9	1.9	1.9	1.8
Nigeria	2.5	2.4	2.4	2.5	2.5	2.5	2.5	2.5
(BRIC countries)								
Brazil	1.7	1.5	1.4	1.1	1.0	0.9	0.9	0.9
China	1.5	1.1	0.8	0.6	0.5	0.5	0.5	0.5
India	2.1	1.9	1.7	1.5	1.5	1.4	1.4	1.4
Russia	0.4	-0.1	0.0	-0.5	-0.3	-0.1	-0.1	-0.1
(selected OECD countries)								
USA	1.1	1.2	1.1	0.9	1.0	0.9	0.9	0.8
UK	0.3	0.3	0.4	0.6	0.6	0.7	0.7	0.7
Japan	0.3	0.4	0.2	0.0	0.0	-0.1	-0.1	-0.1

Source: World Bank WDI & GDF database (accessed on May 4, 2012).

Table 5.3 *GDP, GDP per capita and Their Growth Rate for South Africa, Sub Saharan Africa Region and Selected Countries: 2007 and 2010*

	South Africa		Sub Saharan Africa		Upper Middle Income Countries	
	2007	2010	2007	2010	2007	2010
GDP (constant 2000 US$) (billions)	178.8	187.2	488.3	550.0	6,839.8	7,968.2
GDP per capita (constant 2000 US$)	3,704.8	3,745.6	615.5	644.0	2,847.4	3,249.5
GDP growth (%)	5.6	2.8	6.6	5.1	8.9	7.8
GDP per capita growth (%)	4.4	1.5	4.0	2.6	8.2	7.1

Source: World Bank WDI & GDF database (accessed on May 4, 2012).

Table 5.4 *GDP Growth Rate for South Africa, Sub Sahara Africa Region and Selected Countries: 1965-2010*

	1965	1970	1975	1980	1985	1990	1995	1997	2000	2005	2007	2008	2009	2010
South Africa	8.9	5.2	1.7	6.6	-1.2	-0.3	3.1	2.6	4.2	5.3	5.6	3.6	-1.7	2.8
Sub Saharan Africa	6.4	7.8	1.1	4.0	1.3	1.2	3.8	3.6	3.6	5.7	6.6	5.1	2.0	5.1
(Neighboring middle income countries)														
Angola						-0.3	10.4	7.9	3.0	20.6	22.7	13.8	0.7	5.9
Botwaqna	5.8	17.1	8.4	12.0	7.1	6.8	4.4	10.2	5.9	1.6	4.8	2.9	-4.9	7.2
Ghana	1.4	9.7	-12.4	0.5	5.1	3.3	4.1	4.2	3.7	5.9	6.5	8.4	4.0	7.7
Namibia					0.5	2.5	4.1	4.2	3.5	2.5	5.4	4.3	-0.7	4.8
Nigeria	4.9	25.0	-5.2	4.2	9.7	8.2	2.5	2.7	5.4	5.4	6.4	6.0	7.0	8.7
(BRIC countries)														
Brazil	3.1	8.8	5.2	9.1	7.9	-4.3	4.4	3.4	4.3	3.2	6.1	5.2	-0.6	7.5
China	16.4	19.4	8.7	7.8	13.5	3.8	10.9	9.3	8.4	11.3	14.2	9.6	9.2	10.4
India	21.6	5.2	9.1	6.7	5.2	5.5	7.6	4.1	4.0	9.3	9.8	4.9	9.1	8.8
Russia						-3.0	-4.1	1.4	10.0	6.4	8.5	5.2	-7.8	4.0
(selected OECD countries)														
USA	6.4	0.2	-0.2	-0.3	4.1	1.9	2.5	4.5	4.2	3.1	1.9	0.0	-3.5	3.0
UK	2.8	2.5	-0.6	-2.1	3.6	0.8	3.1	6.2	4.5	2.1	3.5	-1.1	-4.4	2.1
Japan	5.8	4.3	3.1	2.8	5.1	5.2	1.9	1.6	2.9	1.9	2.4	-1.2	-6.3	4.0

Source: World Bank WDI & GDF database (accessed on May 4, 2012).

Table 5.5 *Unemployment Rate for South Africa, Sub Sahara Africa Region and Selected Countries: 1995-2009*

	1995	2000	2005	2007	2009
South Africa	16.9	26.7	26.7	23	23.8
(Neighboring middle income countries)					
Angola					
Botwaqna		15.8	17.6		
Ghana		10.4	3.6		
Namibia		20.3	21.9	37.6	
Nigeria					
(BRIC countries)					
Brazil	6		9.3	8.1	8.3
China	2.9	3.1	4.2	4	4.3
India	2.2	4.3	4.4		
Russia	9.4	10.6	7.2	6.1	8.4
(Selected OECD countries)					
USA	5.6	4	5.1	4.6	9.3
UK	8.6	5.5	4.6	5.2	7.7
Japan	3.2	4.8	4.4	3.9	5

Source: World Bank WDI & GDF database (accessed on May 4, 2012).

Table 5.6 *Total Revenue (% of GDP): 2000-2009*

	2000	2005	2007	2008	2009
South Africa	24.0	26.9	28.9	27.9	25.5
Sub Saharan Africa		17.5	17.8		
(Neighboring middle income countries)					
Ghana	17.2	21.3	13.9	13.9	12.6
Namibia	27.5	25.8	27.3		
Nigeria		0.2	0.2	0.3	
(BRIC countries)					
Brazil	14.0	16.7	16.8	16.7	15.6
China	6.8	8.7	9.9	10.3	10.5
India	9.0	9.9	11.9	10.8	9.7
Russia		16.6	16.6	15.8	13.0
(Selected OECD countries)					
USA		11.2	11.9	10.3	8.5
UK	28.4	27.2	27.6	28.8	26.0
Japan		10.9	10.4	9.2	8.7

Source: World Bank WDI & GDF database (accessed on May 4, 2012).
Notes: The number for Ghana 2000 is actually in 2001. The number for Japan 2005 is actually in 2006. Data for Angola and Botswana was already checked as same as other data tables but the data were not available for these countries.

Table 5.7 *Income Tax Rates in 2007/2008*

Taxable Income (in rand)			Rates of Tax			
R0	-	R112 500			18% of each R1	R112 500
R112 500	-	R180 000	R20 250	+	25% of the amount over	R112 500
R180 000	-	R250 000	R37 125	+	30% of the amount over	R180 000
R250 000	-	R350 000	R58 125	+	35% of the amount over	R250 000
R350 000	-	R450 000	R93 125	+	38% of the amount over	R350 000
R450 000	-	and over	R131 125	+	40% of the amount over	R450 000
Rebates		Primary rebate				R7 740
		Additional rebate (persons of 65 years of age or older)				R4 680
Tax thresholds		The tax thresholds at which liability for normal tax commences, are:				
		• Persons under 65 years of age				R43 000
		• Persons 65 years of age or older				R69 000
Interest exemptions						
		• Persons under 65 years of age				R18 000
		• Persons 65 years of age or older				R26 000

Source: South African Revenue Service (2008).

Table 5.8 *Public Spending on Education (% of GDP) for South African and Selected Countries: 2000-2010*

	2000	2005	2007	2010
South Africa	5.6	5.3	5.2	6.0
Sub Saharan Africa	3.5	4.1		5.0
(Neighboring middle income countries)				
Angola	2.6	2.6		
Botwaqna		9.5	8.0	
Ghana		5.4	5.5	5.5
Namibia	7.9			8.1
Nigeria				
(BRIC countries)				
Brazil	4.0	4.5	5.1	
China				
India	4.4	3.1		
Russia	2.9	3.8		
(Selected OECD countries)				
USA		5.3	5.5	
UK	4.5	5.4	5.5	
Japan	3.7	3.5	3.5	3.8

Source: World Bank WDI & GDF database (accessed on May 4, 2012).
\

Table 5.9 *Public Spending on Education (% of GDP) for South African and Selected Countries: 2000-2010*

	2000	2005	2007	2010
South Africa	18.1	17.9	17.1	19.2
Sub Saharan Africa	15.8		18.9	
(Neighboring middle income countries)				
Angola	6.4	4.4		
Botwaqna		21.5	21.0	
Ghana			24.2	24.4
Namibia	21.0			
Nigeria				
(BRIC countries)				
Brazil	12.0	14.5	16.1	
China				
India	12.7			
Russia	10.6	12.9		
(Selected OECD countries)				
USA		13.7		
UK	11.0	11.8		
Japan	10.5	9.5		

Source: World Bank WDI & GDF database (accessed on May 4, 2012).
Notes: The numbers for Sub Sahara Africa 2000 and 2007 are actually in 1999 and 2008. The numbers for Angola 2000 and 2005 are in 1999 and 2006. The number in 2005 is in 2004.

Table 5.10 *South Africa Structure of Education System, 2008*

BAND	SCHOOL GRADES	NQF LEVEL	QUALIFICATIONS
HIGHER		8	Doctor's degree
		7	Master's degree
			Honours degree
			Postgraduate diploma
		6	General first degree
			Professional first degree postgraduate
			Bachelor's degree
		5	First diploma
			Higher certificate
			Certificate
FURTHER	12	4	Diplomas
	11	3	Certificates
	10	2	
GENERAL	9	1	Grade 9 / Adult Basic Education and Training level 4
	8		
	7		
	6		
	5		
	4		
	3		
	2		
	1		
	R		

Source: UNESCO (2010, August).

Table 5.11 *Primary Education Gross Enrollment Rate for South Africa, Sub Saharan Africa Region and Selected Countries: 1980-2009*

	1980	1985	1990	1995	2000	2005	2007	2009
South Africa	81.3		106.9	117.7	105.9	105.1	104.8	101.7
Sub Saharan Africa					80.5	93.1	96.8	99.9
(Neighboring middle income countries)								
Angola	163.8	83.1	82.2		81.4		114.8	121.5
Botwaqna	88.1	98.9	105.4	104.3	104.0	107.2	107.7	
Ghana	74.2	72.3	69.7	77.8	85.2	90.3	101.4	106.3
Namibia			117.9	124.1	114.8	109.3	109.3	107.5
Nigeria	93.8	104.4	84.2	86.6	97.9	101.8	94.2	83.1
(BRIC countries)								
Brazil	136.7	133.4	140.9		150.7	136.7	127.4	
China	113.7	122.8	128.9	113.2			110.2	111.1
India	83.6	88.4	93.0	96.7	93.8	112.5	113.7	
Russia		104.2	106.8	106.6	103.1	96.6	96.6	98.6

Source: World Bank WDI & GDF database (accessed on May 4, 2012).
Notes: The number in South Africa 1980 is in 1981. The number in Angola 1995 is in 1998.

Table 5.12 *Primary Education Net Enrollment Rate for South Africa, Sub Saharan Africa Region and Selected Countries: 1990-2009*

	1990	1995	2000	2005	2007	2009
South Africa	90.0	95.2	89.7	87.9	87.7	85.1
Sub Saharan Africa			59.2	69.1	72.6	75.1
(Neighboring middle income countries)						
Angola						86.3
Botwaqna	86.9	78.5	80.9	84.5	85.6	
Ghana			64.2	66.5	72.3	76.7
Namibia	79.1	85.7	88.1	87.1	86.6	85.4
Nigeria			64.5	67.2	62.1	
(BRIC countries)						
Brazil			91.7	94.4		94.1
China	97.1	92.5				
India			79.1	89.1	90.3	
Russia		92.2			91.3	93.4

Source: World Bank WDI & GDF database (accessed on May 4, 2012).

Table 5.13 *Secondary Education Gross Enrollment Rate for South Africa, Sub Saharan Africa Region and Selected Countries: 1980-2009*

	1990	1995	2000	2005	2007	2009
South Africa	66.0	79.7	85.3	92.4	95.7	93.8
Sub Saharan Africa			25.0	31.1	33.3	36.0
(Neighboring middle income countries)						
Angola	10.9		14.9		23.8	27.1
Botwaqna	40.1	57.8	74.6	76.9	80.0	
Ghana	35.2		40.5	47.2	53.7	59.1
Namibia	37.9	55.9	60.1	63.1	64.0	
Nigeria	24.3		24.3	34.4	31.5	39.0
(BRIC countries)						
Brazil			104.4	105.8		101.3
China	38.5	51.6	62.1		76.1	80.1
India		45.9	45.3	53.9	57.0	
Russia	94.8	86.8	92.3	83.1	84.7	88.6

Source: World Bank WDI & GDF database (accessed on May 4, 2012).
Notes: The number in South Africa 1995 is in 1994. The number for Brazil 2009 is in 2008. The numbers for Russia 1995 and 2000 are in 1994 and 1999.

Table 5.14 *Secondary Education Net Enrollment Rate for South Africa, Sub Saharan Africa Region and Selected Countries: 1990-2009*

	2000	2005	2007	2009
South Africa	61.9			72.0
Sub Saharan Africa	19.5	23.8	25.8	
(Neighboring middle income countries)				
Angola				
Botwaqna	53.4	55.8	58.8	11.5
Ghana	34.5	40.1	48.4	
Namibia	42.7	49.6	50.7	47.3
Nigeria		27.6	25.8	
Brazil	68.6	78.9		

Source: World Bank WDI & GDF database (accessed on May 4, 2012) for all data except for the number for South Africa 2009 which is obtained from the same database on April 14, 2011).

Table 5.15 *Gini Index*[1]*, 1995-2009*

	1995	2000	2005	2006	2009
South Africa	56.6	57.8		67.4	63.1
(Neighboring middle income countries)					
Angola		58.6			
Botswana	63.0*				
Ghana		40.8*	42.8		
Namibia			63.9*	63.9	
Nigeria	46.5*		42.9*		
(BRIC countries)					
Brazil	60.2	60.1*		56.8	54.7
China	35.7*	39.2*	42.5	42.5	
India			33.4		40.1
Russia	46.1*	39.6*		42.1	
(selected OECD countries)					
USA		40.8			46.8
UK		36.0*	34.0		
Japan			27.8*		37.6*

Sources: WDI & GDF database, US 2009: US Cencus Breau, Japan 2004: Japan Ministry of Internal Affairs and Communications, Statistics Buerau, Japan 2009: CIA World Factbook, UK 2005: CIA World Factbook.
Notes: The numbers with * are in years close to the indicated years. Botswana in 1995 is actually in 1993. Ghana in 2000 is in 1998. Namibia 2005 is in 2004. Nigeria 1995 and 2005 are in 1996 and 2004. Brazil 2000 is in 2001. China 1995 and 2000 are in 1996 and 1999. Russia 1995 and 2000 are in 1996 and 2001. UK 2000 is in 1999. Japan 2005 and 2009 are in 2004 and 2008.

[1] "Gini index measures the extent to which the distribution of income (or, in some cases, consumption expenditure) among individuals or households within an economy deviates from a perfectly equal distribution. A Lorenz curve plots the cumulative percentages of total income received against the cumulative number of recipients, starting with the poorest individual or household. The Gini index measures the area between the Lorenz curve and a hypothetical line of absolute equality, expressed as a percentage of the maximum area under the line. Thus a Gini index of 0 represents perfect equality, while an index of 100 implies perfect inequality" (World Bank WDI & GDF database).

Table 5.16 Poverty Headcount Ratio at National Poverty Line (% of population): 1995-2009

	1995	2000	2005	2006	2009
South Africa	31	38	33	23	
(Neighboring middle income countries)					
Angola					
Botwaqna					
Ghana		39.5		28.5	
Namibia			38		
Nigeria	64.6				
(BRIC countries)					
Brazil	35.1	35.2		26.8	21.4
China	6				
India			37.2		29.8
Russia				11.1	

Sources: World Bank, 2007 (South Africa 2005). WDI & GDF database (all other numbers).
Notes: The number in Sotuh Africa 2005 is in 2004. The number for Ghana 2000 is in 1998. The number for Namibia 2005 is in 2004. The number for Nigeria 1995 is in1996. The number for Brazil 2000 is in 2001. The number for China 1995 is in 1996. The number for India 2009 is in 2010.

Table 5.17 Poverty Headcount Ratio at $2 a day (PPP) (% of population): 1995-2009

	1995	2000	2005	2006	2009
South Africa	39.9	42.9		35.74	31.33
(Neighboring middle income countries)					
Angola		70.21			
Botswaqna					
Ghana		63.34		51.84	
Namibia			51.1		
Nigeria	86.44		83.08		84.49
(BRICs)					
Brazil	20.81	21.71	16.63	14.42	10.82
China	74.13	61.44	36.94		29.79
India			75.62		68.72
Russia	9.01	5.97	1.49	1.22	0.05

Source: World Bank WDI & GDF database.
Notes: The number for Ghana 2000 is for 1998. The number for Namibia 2005 is in 2004. The numbers for Nigeria 1995, 2005, 2010 are in 1996, 2004, and 2009. The number for Brazil 2000 is in 2001. The numbers for China 2000 and 2009 are in 1999 and 2008. The number for India 2009 is in 2010. The numbers for Russia 1995 and 2000 are for 1996 and 2001.

Table 5.18 *Poverty Headcount Ratio at $2 a Day (PPP) (% of population) by Race: 1996 and 2001*

	1996	2001
Black	0.34	0.35
Colored	0.1	0.13
Indian/Asian	0.03	0.03
White	0.01	0.01

Source: Leibbrandt et al. (2006).

Table 5.19 *Unemployment Rate By Race: 1993-2010*

	1993	1997	2001	2005	2006	2007	2008	2009	2010
Black	17	27.1	35.1	31.1			27	27.9	29.5
Colored	15.7	15.2	21.1	22.3			17.3	19.5	22.5
Indian	8.3	9.8	18.7	15.6			6.6	11.3	10.1
White	3.2	3.9	5.6	4.9			10.3	4.6	6.4
Total		22.9	29.5	26.7	25.5	23	22.9	23.6	25.3

Source: PSLSD (1993), OHS (1997), LFS (2001, 2005), NIDS (2008) cited in Leibbrandt et al. (2010). QLFS (2009 & 2010). World Bank WDI & GDF database (accessed on May4, 2012) (Total 1993-2008).

Table 5.20 *Estimated Number of School Leavers Entering the Labor Market for the First-Time Job Seekers: 2002*

	% distribution of South African population	Total number of school-leaving first-time entrants into the labor market	Employment rate of school-leaving first-time entrants into the labour market
Blacks	77	636020	29
Colored	9	74340	50
Indian	3	24780	70
White	11	90860	75
Average	100	826000	37

Source: UNESCO (2011, August).

Table 5.21 *Average Monthly Wage By Race: 1993-2010*

	1993	1997	2001	2005	2007	2008
Black	2,104	2,969	2,588	3,118		2,576
Colored	3,382	3,017	3,834	4,381		3,362
Indian	5,421	5,270	6,315	6,940		7,350
White	10,803	9,508	11,162	12,026		11,240

Source: Leibbrandt, Woolard, McEwen, & Koep (2010).

Table 5.22 *Percentage of the Population Aged 20 and Older, by Race and by Level of Education: 1995 to 2006*

	Year	None	Some form of primary schooling	Grade 7 completed and higher	Total
Black	1995	17.1	21.1	61.9	100
	1997	17.5	19.8	62.7	100
	1998	17.8	20.5	61.7	100
	1999	15	22.4	62.6	100
	2002	14.4	20.4	65.3	100
	2003	13.6	19.3	67.1	100
	2004	13.3	18.7	68	100
	2005	13	18.1	68.9	100
	2006	13.1	16.9	70	100
Coloured	1995	8.8	21.9	69.3	100
	1997	8.7	19	72.3	100
	1998	7.9	18.2	73.9	100
	1999	7.6	20.3	72	100
	2002	7.5	18.5	73.9	100
	2003	7.1	16.8	76.1	100
	2004	6.8	16.5	76.8	100
	2005	5	15	80	100
	2006	4.4	15.9	79.7	100
Indian	1995	5.3	7.8	86.9	100
	1997	5.5	4.6	89.9	100
	1998	4.1	4.7	91.2	100
	1999	3.4	4.7	91.9	100
	2002	3.7	8	88.3	100
	2003	2.9	5.5	91.6	100
	2004	2.4	6.3	91.3	100
	2005	2.6	6.6	90.9	100
	2006	2.9	5.1	92	100
White	1995	0.2	0.6	99.2	100
	1997	0.6	0	99.4	100
	1998	0.5	0	99.5	100
	1999	0.2	0	99.8	100
	2002	0.2	0.4	99.5	100
	2003	0	0	100	100
	2004	0	0	100	100
	2005	0	0	100	100
	2006	0.1	0.6	99.2	100

Sources: Central Statistical Services (undated); and Statistics South Africa (1999, 2000a, 2000b, 2003b, 2004b, 2005, 2006).
Note: Excludes unspecified or “other” educational level. No data is available for 1996, 2000 and 2001.

Table 5.23 *Gross Enrollment Rate in Public Higher Education Institutions by Race: 1986, 1995, 2006*

	1986	1995	2006
African	5%	9%	12%
Colored	9%	10%	12%
Indian	32%	35%	51%
White	61%	61%	59%
Average	11%	14%	16%

Source: South Africa Department of Education, 2009

Table 5.24 *Percentage of Qualified Educators by Race, 1990-2008*

	1990	1994	2005	2006	2007	2008
Black	37	54	90.5	92.1	93	93.9
Coloured	59	71	89.9	92.1	93.2	92.2
Indian	98	93	97.5	97.2	96.6	98.1
White	98	99	99.4	99.5	99.5	99.5
Total	53	64	91.6	92.9	93.8	94.4

Sources: 1990 data from Arnott & Bot, 1993; 1994 data from EduSource Data News No. 10/October 1995; 2005 data from Persal, July 2005; 2006 data from Persal, December 2006; 2007 data from Persal, 2007; and 2008 data from Persal, January 2008.

6. IDENTIFICATION STRATEGIES

This chapter presents the baseline equation and the various methodologies that are used to estimate returns to schooling in the informal and formal sectors.

6.1. Identification Strategies of Return to Education in the Informal and Formal Sectors

6.1.1. Baseline equation

Most empirical studies use the Mincerian approach to estimate rates of return to education. The baseline Mincerian equation used in this dissertation is:

$$(6.1) \quad y_i = \ln Y_i = \alpha + \beta S_i + X_i'\delta + u_i$$

where Y is a measure of income such as earnings; S is a measure of schooling such as years of schooling completed; X is a set of other variables thought to affect income; u is a disturbance representing omitted variables, assumed to be distributed independently of X and possibly of S; and i is an index for a specific individual in the sample (Griliches, 1977; Mincer, 1974). If we use the natural logarithm of hourly wage as $\ln Y$ and years of schooling completed as S, we can estimate the percentage of hourly wage that one additional year of schooling would increase after controlling for background variables X.

It should be noted that the above approach does not take into account the direct and indirect cost of education in calculating rates of return to education. Becker (1962) presented a more comprehensive way to calculate rates of return to education, which included not only the benefits of schooling but also the direct and indirect costs of schooling, all measured through

their present discounted values. Although this is a more comprehensive way to estimate returns to schooling for each level of education, the Mincerian equation has been the most popular method used to estimate returns to schooling because it does not have the data requirements involved in estimating current and future streams of costs and benefits. Also, in this study, returns to education always means “private” returns to schooling, not social returns.

As discussed in the literature review on returns to education in different sectors, the analysis faces the endogeneity of education and sample selection bias from sector choice. Also the heterogeneity of the impact of education on wage is important to factor in.

6.1.2. Dependent, independent, control variables, categorization of sectors/workers, and subgroups of interest

In the rate of return to education regression analysis, the log of “hourly” earnings is used as the dependent variable. The original dataset includes only annual earnings, but this could be divided by number of weeks in a year to calculate weekly earnings. Then hourly earnings could be obtained by dividing weekly earnings by hours of work per week. The wages are before-tax wages, as asked by the survey questionnaire. Although there may be differences in income tax avoidance between workers in the formal and informal sectors, the survey does not provide information on this. However, to test for whether the analysis could change if after-tax earnings were calculated, the empirical analysis below also was carried out on the assumption that only formal workers paid taxes, and the appropriate tax rates were deducted from their survey-based earnings. No significant differences in results were obtained.

The dissertation basically used two types of education-independent variables: years of schooling and a set of dummy variables by level of education. It also employed methods to account for the heterogeneity of returns to schooling, which will be explained in a later section.

The basic set of control variables are experience, experience squared, female, race, marital status, permanent work status,[1] and union participation. Models with more variables were also used, when sample size allowed it, such as language, type of industry, type of occupations, and district dummies.

The dissertation mainly used the informal sector variable already defined by the survey. This definition follows what was called earlier the legalistic definition, which defines someone as working in the informal sector if the employer (institution, business or private individual) is not officially (legally) registered to perform the activity. The informal sector here includes domestic workers as well. Domestic workers are those who are paid to do domestic work, such as housekeepers, nannies, cooks, and gardeners. Note that the research also used different definitions of the informal sector (employment) for a robustness check on the results.

The dissertation divided the formal sector into the public sector and the formal private sector. The main focus is on the comparison of the formal and informal private sector, but the estimations of the public sector are also reported to compare with other sectors. Both wage earners and the self-employed were included in the analysis. Although the analysis couild have been carried out separately for wage earners and the self-employed, sample sizes were greatly reduced when this split was carried out and, instead, the dissertation included all in the same equation, with dummy variables added for the self-employed and head of company (and for any interaction effects) to control for the differences among wage earners and the self-employed in

---

[1]The available categories of work status in the dataset are permanent, a fixed period contract, temporary, casual, and seasonal. However, the study used only the permanent dummy variable: permanent=1 if a worker has a permanent job, and 0 if a worker has a fixed period contract, temporary, casual or seasonal. This is because the estimation of the return to schooling is not very different between the model with only the permanent dummy and the model with all work status dummies, and putting in a lot of independent variables brought about a comformability error in smaller sub-samples by race and gender.

the analysis. In addition, the dissertation closely looked at the impact by race and gender since inequality by race, ethnicity, and gender is a major issue in South Africa.

6.1.3. Dealing with the endogeneity of education and sector sample selection bias

6.1.3.1. Two-stage least squares (2SLS)

This dissertation employed two-stage least squares to control for the endogeneity of education. The model is estimated separately for each sector j.

The first stage equation is

(6.2) $S_{ij} = \gamma Z_i + X_i'\eta + e_i$

where S is schooling, Z is a instrument (or instruments), X is the set of covariates, and e is error term. Then, the second state equation will be

(6.3) $\ln Y_{ij} = \alpha + \beta_1 S_{ij} + X_{ij}'\delta + u_{ij}$

where S is schooling, X is a set of covariates, and u is error term. The instrumental variable affects schooling choices, but is uncorrelated with the ability or error term in the wage equation: $E(\eta_i \mid Z_i) = 0$ (Card, 1999).

It is usually difficult to find a good instrument (or instruments) in datasets from developing countries. Valid instrumental variables often used in the developed countries, where rich datasets are available, are quarter of birth, change in education law, distance to school, parents' education or other family background (Card, 1999). However, it is not easy to obtain these data from surveys in many developing countries. The most popular instruments used for studies on developing countries are parents' education because only that is available (e.g., Barro & Lee, 2010).

In the South Africa Labor Force Survey datasets, none of the above variables are

available. Even father's education and mother's education are not available; there is no direct information on parents' education. Also, information on the relationships between family members except for spouse relationship is not available. Thus, it is impossible to construct parents' education variables using the information in the dataset. The only possible instrumental variable from the dataset is spouse's education (years of schooling). The dataset has information on years of schooling in each household member and information on spouse (household member number is assigned if an individual has a spouse living together inside the household). Based on the information, a spouse's education variable can be constructed. It seems spouse's education is a valid instrument, according to Trostel, Walker, and Woolley (2002). The intuition behind this variable is that years of schooling of husbands as well as of wives are positively correlated. Individuals with comparable levels of education tend to get together as they are likely to share similar experiences and/or similar interests. This hypothesis is called assortative mating and is supported by a number of empirical works (see, for instance, Pencavel, 1998). On the other hand, it is not obvious that spouse's education directly affects individual's earnings.

6.1.3.2. Two-step selection model

As pointed out earlier, we have both endogeneity of education and sector sample selection bias in the analysis of returns to education in the informal private, formal private, and public sectors. However, it is usually difficult to control for both of them at the same time. Most studies actually control for only sector sample selection bias. This section first discusses how to control for sample selection bias and later how to control for both endogenous variables and sample selection bias.

This dissertation uses a two-step selection model to control for sample selection into more the sectors being considered. Lee (1983) developed this two-step model from the Heckman

binary two-step model (1979) to correct for selection bias for more than two groups. The study at first used the following Lee two-step model, based on Trost and Lee (1984).

The wage in the sector j is given by

(6.4) $W_{ij} = \gamma'_j Z_i + \varepsilon_{ji}, \qquad j = 1, 2, ..., M$

where i refers to the i^{th} individual., Z_i is a vector of exogenous variables, and ε_{ji} is a disturbance term.

In the first stage, one estimates a polychotomous model of the choice of employment sector using the logit maximum likelihood method and generate the following selection term for the alternative j.

(6.5) $\hat{\lambda} = \varphi\left[J\left(\hat{\alpha}_j' z_j\right)\right] \Big/ F\left(\hat{\alpha}_j' z_j\right)$

where φ is the standard normal density function, F is the logistic marginal distribution, J is $\Phi^{-1}F$, and Φ denotes the standard normal distribution function. In the second stage $\hat{\lambda}_j$, which is called inverse Mill's ratio, is introduced into the earnings function with other explanatory variables and we can estimate the following equations by OLS.

(6.6) $\ln w_j = \beta_j' x + \sigma_j \hat{\lambda}_j + \upsilon_j$ for j=1,2,...,M

Employing this method, the dissertation could deal with selection bias for more than two sectors, such as when individual workers select the informal private, formal private, and public sectors. For more detail on this method, see Lee (1983) and Trost and Lee (1984).

Lee's method has been criticized because it relies on a strong assumption about the joint distribution of error terms of the equations of interest (Bourguignon et al., 2007; Dahl, 2002; Vijverberg, 1993). Bourguignon et al. (2007) argue that Lee's method could be adopted for very

small samples. Therefore, alternative methods such as those of Dubin and McFadden (1984) or Dahl (2002) were used as well in this study. Bourguignon et al. (2007) also suggest that Dubin and McFadden variant 1 (DMF variant 1), Dubin and McFadden variant 2 (DMF variant 2), and the second model of Dahl (Dahl 2) perform better in general. The DMF variant 1 and variant 2 corrections were developed by Bourguignon et al. (2007) based on the original DMF selection correction. In the present study, Dahl 2 seems to cause multicoliniearities between the selection terms and explanatory variables in the 2SLS due to the large number of selection terms used in the relatively small samples. Also, Bourguignon et al. (2007) state DMF variant 2, which allows for normal error terms in the outcome function, is less robust than original DMF and DMF variant 1. Therefore, the variant one of the Dubin and McFadden selection term developed by Bourguignon et al. (2007) was mainly used in this study.[2]

Another potential problem pointed out by Kuepie et al. (2009) is that the multinomial logit may suffer from the independence of irrelevant alternatives assumptions (IIA). Hausman-type tests for each sector are also employed when the analysis is conducted to examine IIA. Finally, in order to strengthen the methodology, instrumental variables are needed in the first stage equation. In general, exclusion restrictions need to be held to make the estimates credible, which means instruments are required. This study used spouse formal private sector dummy, spouse informal private sector dummy, family formal private sector dummy, and family informal private sector dummy as instruments.

Utilizing this method, the dissertation used the following groups in the sample selection analysis: 1) working in the public sector, 2) working in the formal private sector, 3) working in

---

[2]See Bourguignon et al. (2007) for details of the formula for the selection terms of DMF, DMF variant 1, DMF variant 2, Dahl 1, and Dahl 2.

the informal private sector, and 4) unemployed or out of labor force.

As shown in the literature review, most previous studies used the Heckman or Lee type two-step selection methods to control for sectoral allocation. This method was a major tool to control for selection bias of sectoral allocation in this dissertation as well.

The basic set of explanatory variables used for the choice model include: level of education dummies (primary, lower secondary, upper secondary, certificate with grade 12 and diploma with grade 12, and bachelor's degree and above), age, female, races, marital status, and dependent ratio. Also, spouse public sector, spouse formal private sector, household public sector, and household formal private sector dummies were included as the "instruments" which were not in the wage equation and likely to affect sector choice but did not directly affect wage.

6.1.3.3. Two-stage least squares with two-step selection correction

The two-stage least squares and two-step selection correction models correct only for the endogeneity of education or sector sample selection bias, respectively. Therefore, this study combined these two methods to correct for both of them at the same time. Wooldridge (2002b, p. 567)[3] actually suggests a similar method when there are both endogeneity and sample selection issues, using the Heckman selection correction for two categories. Wooldridge (2002b) suggests that this procedure gives a consistent estimate of the coefficients. The method should be applied to more than two categories using the two-step selection correction for multiple choices.

In this dissertation, the first stage equation was

(6.7) $S_{ij} = \gamma Z_{ij} + X_{ij}'\eta + e_i$

where S is schooling, Z is instrument, X is the set of covariates, and e is error term.

[3]Rivers and Vuong (1988) actually use the methods.

j stands for a certain sector. The choice equation was

(6.8) $\Pr(J_i = 1 | W_i, X_i) = \Phi(\rho_j W_i + X_i \delta_j)$

where J is being in a certain sector, W is instrument, X is a set of covariates, and ω is an error term. The equation 7.6 is estimated by the multinomial logit covering four categories (public, formal private, informal private, and unemployed/out of labor force). Based on the estimation of the choice model, the Lee's inverse Mill's ratio, Dubin and McFadden selection terms, and Dahl selection terms, and variants of Dubin and McFadden selection terms were calculated.

Then the second stage equation for sector *j* was

(6.9) $\ln Y_{ij} = \alpha_j + \beta_j S_{ij} + X_{ij}' \delta_j + \gamma_j \hat{\lambda}_{ij} + u_{ij}$

where S is schooling, X is a set of covariates, $\hat{\lambda}_j$ is inverse Mill's ratio or other selection terms, and u is error term.

6.1.3.4. Blinder-Oaxaca decomposition method

A popular method of comparing wage differentials between groups is the wage decomposition methodology of Oaxaca (1973) and Blinder (1973). Some studies researching the topic of this dissertation, such as Lassible and Tan (2005) and Pagán and Tijerina-Guajardo (2000), used the wage decomposition method. This dissertation used the Blinder-Oaxaca decomposition for the 2SLS model that controls for the endogeneity of education with the Heckman selectivity correction.

First, the basic Blinder-Oaxaca decomposition method developed by Oaxaca (1973) and Blinder (1973) is discussed, using two categories: the formal and informal sectors. Let $\ln(\bar{w}_F)$ and $\ln(\bar{w}_I)$ be the means of the (natural) logs of wages in the formal (F) and informal sectors

(I).[4] When the standard log wage model is estimated separately for workers in the formal sector and in the informal sector, the following equation holds since regression lines pass through the means of the variables.

(6.10) $\ln(\bar{w}_F) - ln(\bar{w}_I) = \bar{X}_F' b_F - \bar{X}_I' b_I$

where $\bar{X}_F$ and $\bar{X}_I$ are vectors containing the means of the variables for males and females, respectively, and b_F and b_I are the estimated coefficients.

Given these results, the log wage differential can be decomposed in two ways. Letting $\Delta\bar{X}' = \bar{X}_F' - \bar{X}_I'$ and $\Delta b = b_F - b_I$, (6.10) can be written as

(6.11) $\ln(\bar{w}_F) - ln(\bar{w}_I) = \Delta\bar{X}' b_F + \bar{X}_I' \Delta b$

or

(6.12) $\ln(\bar{w}_F) - ln(\bar{w}_I) = \Delta\bar{X}' b_I + \bar{X}_F' \Delta b$

The first term of either Eq. (6.11) or Eq. (6.12) is the part of the log wage differential due to different (average) characteristics of workers in the formal and informal sectors, and the second term is the part of the differential due to different coefficients or different wage structures. If workers in two sectors receive identical returns for the same characteristics in the absence of discrimination, this second term can be interpreted as the part of the log wage differential due to discrimination. A critical assumption here is that labor supply and individual characteristics are fixed, and will not respond to changes in wages caused by the elimination of discrimination (Butler, 1982).

To control for selection bias, Newman and Oaxaca (2004) and others (e.g., Idson &

[4] We could also compare the public and formal private sector or the public and informal private sector but here the comparison of the formal and informal sector is used to explain the method.

Feaster, 1990) employed the Heckman or Lee two-step model, which estimates the probability of falling in a certain category.

If the dissertation employed this method, the wage decomposition equation would be

(6.13) $\ln(\bar{w}_F) - ln(\bar{w}_I) = \Delta\bar{X}'b_F + \bar{X}_I'\Delta b + (\theta_F\bar{\lambda}_F - \theta_I\bar{\lambda}_I)$

or

(6.14) $\ln(\bar{w}_F) - ln(\bar{w}_I) = \Delta\bar{X}'b_I + \bar{X}_F'\Delta b + (\theta_F\bar{\lambda}_I - \theta_F\bar{\lambda}_I)$

As well, the original wage equations are estimated by two-stage least squares to control for the endogeneity of education.

Another issue to deal with concerns computing wage decomposition. One has to choose either Eq. (6.11)/Eq. (6.13) or Eq. (6.12)/Eq. (6.14) in the empirical work, and either of the two equations could lead to a substantial difference in results. For example, in Oaxaca's (1973) study which used data on Whites, Eq. (6.12) estimated that 52.9% of the male-female log wage differential was due to discrimination, while Eq. (6.13) yielded an estimate of 63.9%. A much larger difference occurred in Ferber and Green (1982), in a study of pay discrimination for a sample of university professors, where Eq. (6.12) yielded an estimate of 2% of the wage differential due to discrimination, while Eq. (6.13) gave an estimate of 70%. To tackle this problem, Reimers (1983) and Cotton (1988) used weighted averages.

Reimers (1983) developed the Oaxaca wage decomposition equation into the following equation:

(6.15) $\ln(\bar{w}_F) - \ln(\bar{w}_F) = \Delta X'[Db_F + (I - D)b_I] + [\bar{X}_F(I - D) + \bar{X}_I D]\Delta b$

$$+(\theta_F\bar{\lambda}_F - \theta_I\bar{\lambda}_I)$$

where I is the identity matrix and D is a diagonal matrix of weights. In general., we will

obtain different measures of the impact of discrimination, depending on the choice of weights in matrix D. Reimers claimed that the non-discrimination wage function lies somewhere between them; therefore, he chose D = (0.5)I. By contrast, Cotton (1988) chose the weighting matrix $\Omega_c = l_w I$, where l_w is the fraction of the sample made up by a majority group, claiming that the nondiscriminatory structure should be more similar to the structure that holds for the larger group. Oaxaca and Ransom (1994) pointed out that neither of these solutions is completely satisfactory since each researcher cited above chose the weight arbitrarily. They proposed another weighting method, which is

(6.16) $\Omega_o = (X^{'} X)^{-1} (X_F^{'} X_F)$

where X is the observations matrix for the pooled sample and X_F is the observation matrix for the majority sample (for the present study, it was workers in the formal sector). Another alternative is three-fold decomposition which uses the same reference group for both effects, but adds a third interaction term

(6.17) $I = \{E(X_F) - E(X_I)\}'(\beta_F - \beta_I)$

(Jann, 2008). The three-fold decomposition was used for the analysis in the dissertation.

6.1.4. Dealing with the heterogeneity of return to schooling

In addition to selection bias, heterogeneity of returns to education is important to consider. This dissertation employed the quantile regression approach and a piecewise linear spline function that was employed by Kuepie et al. (2009) to consider the heterogeneity of the impact of education on earnings.

6.1.4.1. Quantile regression

One popular way to deal with heterogeneity is by using quantile regression (Koenker &

Bassertt, 1978). One can use quantile regression to estimate earnings and return to schooling gaps between formal and informal workers at different points along the conditional earnings distribution. As least squares models the mean of the distribution of the dependent variable Y conditional on regressors Z, quantile regressions give models for different percentiles of this distribution. The τ-th quantile of Y conditional on Z is given by:

(6.17) $Q_\tau(Y_i \mid Z_i) = Z_i'\phi(\tau)$

where coefficient $\phi(\tau)$ is the slope of the quantile line giving the effect of changes in Z on the τ-th conditional quantile of Y. Estimation for different quantiles (τ from 0 to 1) produces regression lines for the various percentiles of the conditional distribution of Y such that at least a τ proportion of the regression residuals are below the estimated regression line and approximately a $(1-\tau)$ fraction is above it. For example, median regression $(\tau = 0.5)$ splits the sample in half (half of the residuals are above and half are below the regression line) and gives the same results as ordinary least squares when the distribution is symmetric. Using quantile regression, this dissertation computed the years of schooling coefficients at different quantiles for each sector.

According to Koenker and Bassertt (1978), quantile regression is defined as the minimization problem shown below. Letting $\{x_t : 1,...,T\}$ denote a sequence of (row) K-vectors of a known design matrix, suppose $\{y_t : t = 1,...,T\}$ is a random sample on the regression process $u_t = y_t - x_t\beta$ having distribution function F. The *θth regression quantile*, $0 < \theta < 1$, is defined as any solution to the minimization problem:

$$\min_{b \in \mathbb{R}^K} \left[\sum_{t \in \{t: y_t \geq x_t b\}} \theta \mid y_t - b \mid + \sum_{t \in \{t: y_t < x_t b\}} (1-\theta) \mid y_t - x_t b \mid \right]$$

In the location model ($K = 1, x_t = 1$, for all t), the two minimization problems coincide. The least absolute error estimator is the regression median, i.e., the regression quantile for θ =1/2.

For multiple-choice sample selection models, one can use a similar approach proposed by Fitzenberger (2003) for multiple-choice sample selection models and quantile regression. The probability of participation in different sectors is determined in the first stage using the multinomial logit model. In the second-stage quantile wage equations, one then puts the DMF variant 1 selection correction term to correct for selectivity.

In addition, this study conducted the instrumental variable quantile regression with the DMF variant 1 selection correction to correct for both the endogeneity of schooling and sample selection bias. The instrumental variable quantile regression was developed by Chernozhukov and Hansen (2004).

6.1.4.2. Piecewise linear spline function

This dissertation used a piecewise linear spline function employed by Kuepie et al. (2009) as well. The piecewise linear splice function allows the relationship between schooling and earnings to vary across different levels of education. More specifically, the return to schooling is allowed to be different at four levels of education: primary, lower secondary, upper secondary, and higher education. The education variables introduced have the form $s_k(e)$ with e as the years of completed schooling in the k levels (k :1...4):

(6.19) $s_1(e) = \begin{cases} e, & e \le 6 \\ 6, & \text{e} > 6 \end{cases}$ $s_2(e) = \begin{cases} 0, & e \le 6, \\ e-6, & 6 < e \le 10, \\ 4, & e > 10. \end{cases}$

$s_3(e) = \begin{cases} 0, & e \le 10, \\ e-10, & 10 < 3 \le 13, \\ 3, & e > 13. \end{cases}$ $s_4(e) = \begin{cases} 0, & e \le 13, \\ e-13, & e > 13. \end{cases}$

6.2. Identification Strategies of the Estimation of the Effect of Schooling on Informal and Formal Sector Employment

In this section, the empirical strategies used to estimate the effect of schooling on the likelihood of working in formal and informal employment are discussed.

6.2.1. Baseline equation

To analyze the sectoral allocation or choice between sectors, either the probit or the binary logit model is used when there are only two categories, while the multinomial logit model is used with more than two categories. The baseline probit model is:

(6.20) $\Pr(I_{it} = 1 \mid X_{it}) = \Phi(X_{it}^{'}\beta_t)$

where Φ is the cumulative density function of a standard normal random variable. X_{it} is a vector of explanatory variables affecting the employment decision, including education variable(s), and β_{it} is a parameter vector of the alternative chosen. t is either 0 or 1; one number often refers to the formal sector, while the other refers to the informal sector.

The baseline equation for binary logit and multinomial logit model is:

(6.21) $P_j = \dfrac{\exp(\alpha_j^{'} X)}{1 + \sum_{j=1}^{N} \exp(\alpha_j^{'} X)}$

where X is a vector of explanatory variables affecting the employment decision, including

education variable(s), and α_j is a parameter vector of the alternative chosen. j refers to each type of employment. In a binary logit, j =0 or 1, and one number usually refers to the formal sector while the other refers to the informal sector. With a multinomial logit, there are more than two categories to study. In this study, j =0 is unemployed or out of the labor force (an omitted category), j =1 is working in the public sector, j =2 is working in the formal private sector, and j =3 is working in the informal private sector. This study used the probit and multinomial logit models.

6.2.2. Dependent, independent, and control variables

The dependent variable in the probit equation to estimate the effect of schooling on the probability of working in the informal and formal private sectors is a binary variable on whether a worker is in the informal sector or in the formal sector. A multinomial model was also estimated, where the emoployment probability is a discrete variable with more than two categories. As noted earlier, in defining the informal versus the formal sector of employment, this study first used the sector variable defined by the survey and then used other sector variables defined differently by the author.

The key independent variables of interest are the education variables, which were either continuous, years of schooling variable or a set of level of education variables. The covariates included in the basic model are: age, female, races, languages, marital status, dependent ratio, training received, participation in government job creation program, spouse public sector participation, spouse formal private sector participation, household public sector participation, and household formal private sector participation.

6.2.3. Dealing with the endogeneity of education

The analysis of the effects of schooling on the probability of working in the formal and informal sectors is usually conducted as part of the selection bias correction for the analysis of the rate of return to schooling. Since most studies utilize the Heckman or Lee type of selection correction method, these studies have estimated the probability of working in these sectors in order to adjust the wage equations for selectivity bias. Therefore, the probability analysis was secondary and few studies have corrected for the endogeneity of schooling in estimating the probabilities of employment among sectors.

This dissertation estimated the multinomial logit and probit models without controlling for the endogeneity of schooling first, and then the instrumental variable probit model was estimated, controlling for the endogeneity of schooling.

7. DESCRIPTIVE STATISTICS OF THE SAMPLES

This chapter provides descriptive statistics for variables used in the analysis. The summary statistics on workers in different sectors are shown in Tables 7.1-7.9.

Table 7.1 reports the number of workers by sector, gender, and race. It shows that the majority of workers belong to the formal private sector across race and gender except for Black females. However, the proportion of workers in the informal sectors is not small, especially for Blacks and Coloreds. The proportion of workers in the informal private sector is smaller than the one in the formal private sector, but larger than the one in the public sector in general, except for Whites and Asians. More workers work in the public sector than in the informal private sector for Whites and Asians.

The proportion of workers who are employed in the public sector ranges from 14.3% to 21.5% across race and gender groups. Females are more likely to work in the public sector compared to males, while the trend is opposite for the private formal sector. The share of workers in the public sector ranges from 16.5% to 21.5% across race groups for females, while the share of workers in the public sector ranges from 14.3% to 16.6% for males. The proportion of workers in the formal private sector vary are from 58.2% to 78.9% for males, while the proportions of workers in the formal private sector are from 35.5% to 77.5% for females. Differences in the participation on the informal sector by race and gender are mixed. While males are more likely to work in the informal private sector than Whites and Asians, females are much more likely to work in the informal private sector than males for Blacks. The proportion of workers in the informal private sector for males is 5.8% for White and 19.4% for Asian, compared to 4.2% for White females and 5.1% for Asian females.

Table 7.2 is slightly different from Table 7.1. It displays the number of workers by sector, gender, and race with the proportions of workers by race in each sector. We can see that the majority of workers in the informal private sector are Black, while only a small number of Whites and Asians work in the informal private sector. The proportion of Colored workers in the informal private sector lies in between Whites/Asians and Blacks. Disproportionately more Blacks work in the informal private sector, while disproportionately more Whites, Colored, and Asians work in the formal private sector. The public sector reflects the proportions of races in the country. The distribution of the types of job for different races are not exactly the same, but at least the numbers in Table 7.2 show that the public sector hires workers from different population groups equally. In contrast, the proportion of Blacks in the formal private sector is a bit lower than the overall proportions of Blacks in the whole labor market. In addition, the proportions of Whites, Asians, and Colored in the formal private sector are a little disproportionately higher than the overall proportions of races in the whole labor market. Also, the proportion of Blacks in the informal private sector is disproportionately high, while the proportions of Whites, Colored, and Asians in the informal sector are disproportionately low.[5]

Column 1 in Table 7.2 shows the proportion of workers by race in the overall labor market: 10.1% of workers in the labor market are White, 70% are Black, 16.9% are Colored, 2.7% are Asian, and 0.2% are of other races. These figures differ by sector. In the public sector, 10.3% of workers are White, 71.1% are Black, 16.2% are Colored, 2.3% are Asian. In the formal private sector, 14.2% are White, 61.3% are Black, 20.4% are Colored, and 3.8% are Asian; in the informal private sector, 1.9% are White, 86.3% are Black, 10.7% are Colored, and 0.8% is Asian.

[5]Proportions of each race in each sector are also shown in Table 7.3.

These differences are similar for the male and female samples.

The descriptive statistics reported in Table 7.3 show the sample means for various variables by sector and gender. Workers in the public sector are characterized by the highest wage, highest years of education completed, most experienced, most training experience, and highest rate of trade union participation. The situation of workers in the formal private sector is the second best in general. The compensation and qualifications of workers in the informal private sector are the worst. Finally, the compensation and qualifications of males are better than females in general, except that years of schooling for females are higher than for males in the public and formal private sectors.

Let us look at the statistics in more detail. The average hourly wage for the public sector is 30.97, the highest among all three sectors. The average hourly wages for the formal private sector is 15.53 rand, about half of that for the public sector. The figure for the informal private sector is half of that for the formal private sector, which is 8.37 rand. The difference in average hourly wages between sectors is smaller for males, but the differences for males and females are similar. The average hourly wages for females is lower than that for males. The average years of schooling is also the highest for the public sector, and the lowest for the informal private sector for total, male and female. Average years of schooling are equal to 11.63 for the public sector, 9.48 for the formal private sector, and 6.90 for the informal private sector for total. The average of years of schooling for the public and formal private sectors is approximately one year higher for females than males.

The ratio of males to females is about 50:50 for the public sector. In contrast, the proportion of female is much smaller in the formal private sector, 36%. The proportion of females in the informal sector is 58%, slightly greater than males.

The proportion of married people is the largest in the public sector for both males and females. The proportions of married workers for male and female in the public sector are 69% and 50%. Male informal private sector workers have the lowest proportion of married people, equal to 49%. The proportion of female married workers in the formal and informal private sectors is similar and smaller than in the public sector (45% and 46%, respectively).

Looking at the work status of workers across sectors, the public sector has the largest proportion of workers with permanent status: 88% for males and 83% for females. The proportion of permanent workers in the formal private sector is a bit lower than in the public sector: 72% for males and 68% for females. The proportion of permanent workers in the informal private sector is much lower than in the other sectors: 28% for males and 48% for females. As one can expect, the proportion of temporary or casual workers is much higher in the informal private sector than in the other sectors. The proportions of temporary workers for males and females are 37% and 30% in the informal sector. In contrast, they are 12% for both males and females in the formal private sector and 7% and 5% for males and females in the public sector. The proportions of casual workers in the informal private sector are 29% for males and 18% for females, and 7% for male and 9% for females in the formal private sector. The proportion of casual workers in the public sector is 2% for both male and female.

Workers in the public sector are more experienced, followed by workers in the formal private sector. The years of experience in the current main occupation are 11.44 years for males and 9.73 for females in the public sector. The years of experience in the current main occupation in the formal private and informal private sectors are 6.37 for males and 5.13 for females, and 3.55 for males and 4.98 for females, respectively.

Also, a greater proportion of workers receive training in the public sector than in the

other sectors, while the proportion of workers with training experience is the lowest in the informal private sector. The proportion of workers who receive some kind of training in the public sector is equal to17% and 15% for males and females, respectively. They are 14% for males and 12% for females in the formal private sector, while they are 8% for males and 6% for females in the informal private sector.

The union participation rate is very high in the public sector, while it is much lower in the formal private sector and remarkably low in the informal private sector. The union participation rates are 74% and 69% for males and females in the public sector, 30% for males and 25% for females in the formal private sector, and 2% for both males and females in the informal private sector.

Tables 7.4 to 7.7 show descriptive statistics by race. The differences across variables are very similar to those in Table 7.3, but the compensation and qualifications for Whites are much higher than for other population groups. Asians are less well off than Whites, but still much better than Blacks and Colored. Compensation and qualifications for Blacks and Colored are much worse than for Whites and Asians, but the worst are for the Colored population. For example, average hourly wages in rand in the public, formal private, and informal private sectors are, respectively, 51.25, 49.99, and 42.85 for Whites, while they are 29.30, 12.21, and 7.69 for Blacks. The average years of schooling in the public, formal private, and informal private sectors are, respectively, 13.02, 12.50, and 11.51 for Whites, while they are 11.53, 8.92, and 6.78 for Blacks.

The number of workers by industry, sector, and gender is displayed in Table 7.8. In addition, Table 7.9 displays the number of workers by occupation, sector, and gender.

Figures 7.1 and 7.2 show the wage distribution and the distribution of years of schooling

by sector, respectively. In Figure 7.1, all distributions are close to normal distribution, but the means are different. The mean of the wage in the public sector is the largest, as expected, followed by the formal private and then the informal private. Figure 7.2 indicates many zero schooling observations (the largest number across different years of schooling) in the informal sector. The distribution for the public and formal private is skewed toward the right.

Figure 7.3 displays years of schooling vs. log hourly wage by sector. These figures imply that years of schooling and log hourly wage are positively correlated in all sectors. The correlation is the strongest for the public sector, followed by the formal public sector. The correlation is the smallest for the informal sector.

Finally, as noted before, 28.3% of workers did not report their earnings. However, Figure 7.4 shows that the distributions of education variables between those who report earnings and those who did not report earnings are not substantially different. Thus, the study conducted the return to education analysis for workers who reported their earnings.

Table 7.1 *Number of Workers by Sector, Gender, and Race, Part I*

	Total				Male				Female			
	All sectors	Public	Private, formal	Private, informal	All sectors	Public	Private, formal	Private, informal	All sectors	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Total	25,474	4,495	13,906	7,073	14,024	2,200	8,838	2,986	11,450	2,295	5,068	4,087
	100.0%	17.6%	54.6%	27.8%	100.0%	15.7%	63.0%	21.3%	100.0%	20.0%	44.3%	35.7%
White	2,575	463	1,981	131	1,475	226	1,164	85	1,100	237	817	46
	100.0%	18.0%	76.9%	5.1%	100.0%	15.3%	78.9%	5.8%	100.0%	21.5%	74.3%	4.2%
Black	17,824	3,194	8,526	6,104	9,673	1,512	5,633	2,528	8,151	1,682	2,893	3,576
	100.0%	17.9%	47.8%	34.2%	100.0%	15.6%	58.2%	26.1%	100.0%	20.6%	35.5%	43.9%
Colored	4,317	727	2,835	755	2,369	393	1,667	309	1,948	334	1,168	446
	100.0%	16.8%	65.7%	17.5%	100.0%	16.6%	70.4%	13.0%	100.0%	17.1%	60.0%	22.9%
Asian	698	105	535	58	467	67	356	44	231	38	179	14
	100.0%	15.0%	76.6%	8.3%	100.0%	14.3%	76.2%	9.4%	100.0%	16.5%	77.5%	6.1%
Other	60	6	29	25	40	2	18	20	20	4	11	5
	100.0%	10.0%	48.3%	41.7%	100.0%	5.0%	45.0%	50.0%	100.0%	20.0%	55.0%	25.0%

Note: Proportions shown are the proportions of workers in each sector by race.

Table 7.2 *Number of Workers by Sector, Gender, and Race Part II*

	Total				Male				Female			
	All sectors	Public	Private, formal	Private, informal	All sectors	Public	Private, formal	Private, informal	All sectors	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Total	25,474	4,495	13,906	7,073	14,024	2,200	8,838	2,986	11,450	2,295	5,068	4,087
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
White	2,575	463	1,981	131	1,475	226	1,164	85	1,100	237	817	46
	10.1%	10.3%	14.2%	1.9%	10.5%	10.3%	13.2%	2.8%	9.6%	10.3%	16.1%	1.1%
Black	17,824	3,194	8,526	6,104	9,673	1,512	5,633	2,528	8,151	1,682	2,893	3,576
	70.0%	71.1%	61.3%	86.3%	69.0%	68.7%	63.7%	84.7%	71.2%	73.3%	57.1%	87.5%
Colored	4,317	727	2,835	755	2,369	393	1,667	309	1,948	334	1,168	446
	16.9%	16.2%	20.4%	10.7%	16.9%	17.9%	18.9%	10.3%	17.0%	14.6%	23.0%	10.9%
Asian	698	105	535	58	467	67	356	44	231	38	179	14
	2.7%	2.3%	3.8%	0.8%	3.3%	3.0%	4.0%	1.5%	2.0%	1.7%	3.5%	0.3%
Other	60	6	29	25	40	2	18	20	20	4	11	5
	0.2%	0.1%	0.2%	0.4%	0.3%	0.1%	0.2%	0.7%	0.2%	0.2%	0.2%	0.1%

Note: The proportions shown are the proportions of workers by race in each sector

Table 7.3 *Descriptive Statistics, Total*

Variable	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Hourly wage	30.97	15.53	8.37	31.35	16.65	10.53	30.24	13.51	6.88
	(28.35)	(29.05)	(19.34)	(27.58)	(28.42)	(25.58)	(29.10)	(30.06)	(13.20)
Years of schooling	11.63	9.48	6.90	11.05	9.17	7.19	12.09	10.03	6.69
	(3.67)	(3.68)	(3.95)	(3.89)	(3.80)	(3.91)	(3.47)	(3.39)	(3.96)
Age	40.64	36.85	40.08	40.88	37.39	38.30	40.42	35.91	41.38
	(9.98)	(11.13)	(12.05)	(10.12)	(11.39)	(12.89)	(9.84)	(10.60)	(11.22)
Female	0.51	0.36	0.58						
	(0.50)	(0.481)	(0.494)						
White	0.11	0.14	0.02	0.10	0.13	0.03	0.10	0.16	0.01
	(0.31)	(0.350)	(0.135)	(0.305)	(0.34)	(0.17)	(0.30)	(0.37)	(0.11)
Black	0.71	0.61	0.87	0.68	0.64	0.85	0.73	0.57	0.88
	(0.46)	(0.49)	(0.34)	(0.47)	(0.48)	(0.35)	(0.44)	(0.49)	(0.33)
Colored	0.16	0.20	0.11	0.18	0.19	0.10	0.15	0.23	0.11
	(0.37)	(0.40)	(0.31)	(0.39)	(0.39)	(0.31)	(0.35)	(0.42)	(0.31)
Asian	0.02	0.04	0.01	0.03	0.04	0.01	0.02	0.04	0.00
	(0.15)	(0.19)	(0.09)	(0.17)	(0.20)	(0.12)	(0.13)	(0.18)	(0.06)
Married	0.59	0.54	0.47	0.69	0.59	0.49	0.50	0.45	0.46
	(0.49)	(0.50)	(0.50)	(0.46)	(0.49)	(0.50)	(0.50)	(0.50)	(0.50)

Variable	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Permanent	0.86	0.71	0.40	0.88	0.72	0.28	0.83	0.68	0.48
	(0.35)	(0.46)	(0.49)	(0.32)	(0.45)	(0.45)	(0.37)	(0.47)	(0.50)
Fixed	0.06	0.07	0.04	0.05	0.07	0.04	0.06	0.07	0.03
	(0.23)	(0.25)	(0.19)	(0.22)	(0.25)	(0.20)	(0.24)	(0.25)	(0.18)
Temporary	0.07	0.12	0.33	0.05	0.12	0.37	0.09	0.12	0.30
	(0.25)	(0.32)	(0.47)	(0.21)	(0.32)	(0.48)	(0.28)	(0.32)	(0.46)
Casual	0.02	0.08	0.22	0.02	0.07	0.29	0.02	0.09	0.18
	(0.13)	(0.271)	(0.42)	(0.13)	(0.26)	(0.45)	(0.13)	(0.29)	(0.38)
Experience in the current main occupation	10.60	5.91	4.43	11.44	6.37	3.55	9.73	5.13	4.98
	(9.39)	(7.49)	(6.36)	(9.63)	(7.98)	(5.54)	(9.03)	(6.48)	(6.78)
Training	0.16	0.13	0.07	0.17	0.14	0.08	0.15	0.12	0.06
	(0.37)	(0.34)	(0.25)	(0.38)	(0.35)	(0.27)	(0.35)	(0.33)	(0.23)
Union	0.72	0.28	0.02	0.74	0.30	0.02	0.69	0.25	0.02
	(0.45)	(0.45)	(0.15)	(0.44)	(0.46)	(0.13)	(0.46)	(0.43)	(0.16)

Note: Standard deviations are in parentheses.

Table 7.4 *Descriptive Statistics, White*

Variable	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Hourly wage	51.25	49.99	42.85	55.32	57.54	42.93	47.27	39.59	42.72
	(44.71)	(68.08)	(53.77)	(49.28)	(72.42)	(46.73)	(39.61)	(60.19)	(65.64)
Years of schooling	13.02	12.50	11.51	12.69	12.47	11.27	13.27	12.55	11.93
	(2.11)	(2.19)	(2.72)	(2.29)	(2.29)	(2.38)	(1.82)	(2.04)	(3.23)
Age	43.50	40.80	43.58	43.25	41.74	43.87	43.71	39.48	43.04
	(9.77)	(11.88)	(11.90)	(9.48)	(12.04)	(12.40)	(9.94)	(11.51)	(11.04)
Female	0.51	0.41	0.35						
	(0.50)	(0.49)	(0.48)						
Married	0.76	0.76	0.65	0.78	0.78	0.59	0.75	0.73	0.76
	(0.43)	(0.43)	(0.48)	(0.41)	(0.41)	(0.50)	(0.44)	(0.44)	(0.43)

Variable	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Permanent	0.95	0.93	0.74	0.96	0.94	0.73	0.94	0.92	0.75
	(0.21)	(0.25)	(0.45)	(0.19)	(0.23)	(0.46)	(0.24)	(0.28)	(0.45)
Fixed	0.03	0.03	0.04	0.02	0.03	0.07	0.03	0.03	0.00
	(0.17)	(0.17)	(0.19)	(0.15)	(0.16)	(0.26)	(0.18)	(0.17)	(0.00)
Temporary	0.02	0.03	0.15	0.01	0.02	0.13	0.03	0.04	0.17
	(0.14)	(0.17)	(0.36)	(0.12)	(0.15)	(0.35)	(0.16)	(0.19)	(0.39)
Casual	0.00	0.01	0.07	0.00	0.01	0.07	0.00	0.02	0.08
	(0.00)	(0.11)	(0.27)	(0.00)	(0.08)	(0.26)	(0.00)	(0.13)	(0.29)
Experience in the	14.14	7.47	6.67	16.06	8.14	7.60	12.43	6.68	5.50
current main occupation	(9.60)	(8.44)	(7.40)	(10.35)	(9.02)	(7.95)	(8.49)	(7.61)	(6.80)
Training	0.18	0.19	0.23	0.21	0.20	0.21	0.16	0.17	0.27
	(0.39)	(0.39)	(0.42)	(0.41)	(0.40)	(0.41)	(0.37)	(0.38)	(0.45)
Union	0.72	0.24	0.11	0.74	0.27	0.13	0.72	0.20	0.08
	(0.45)	(0.43)	(0.32)	(0.44)	(0.44)	(0.35)	(0.45)	(0.40)	(0.29)

Note: Standard deviations are in parentheses.

Table 7.5 *Descriptive Statistics, Black*

Variable	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Hourly wage	29.30	12.21	7.69	28.75	13.17	9.47	29.82	10.35	6.50
	(26.65)	(21.36)	(18.74)	(23.59)	(17.87)	(25.52)	(29.14)	(26.75)	(12.08)
Years of schooling	11.53	8.92	6.78	10.88	8.62	6.98	12.11	9.49	6.64
	(3.91)	(3.70)	(3.95)	(4.11)	(3.77)	(3.90)	(3.62)	(3.51)	(3.99)
Age	40.51	36.21	39.97	40.75	36.63	38.03	40.29	35.39	41.33
	(9.89)	(10.75)	(12.18)	(10.16)	(10.96)	(13.03)	(9.63)	(10.29)	(11.35)
Female	0.53	0.34	0.59						
	(0.50)	(0.47)	(0.49)						
Married	0.55	0.47	0.45	0.66	0.54	0.47	0.46	0.35	0.44
	(0.50)	(0.50)	(0.50)	(0.47)	(0.50)	(0.50)	(0.50)	(0.48)	(0.50)

Variable	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Permanent	0.84	0.65	0.38	0.87	0.68	0.26	0.82	0.60	0.45
	(0.37)	(0.48)	(0.48)	(0.34)	(0.47)	(0.44)	(0.39)	(0.49)	(0.50)
Fixed	0.06	0.08	0.04	0.05	0.08	0.04	0.06	0.09	0.04
	(0.23)	(0.27)	(0.19)	(0.23)	(0.27)	(0.20)	(0.24)	(0.28)	(0.19)
Temporary	0.08	0.16	0.37	0.06	0.15	0.43	0.10	0.17	0.34
	(0.27)	(0.36)	(0.48)	(0.23)	(0.36)	(0.49)	(0.30)	(0.38)	(0.47)
Casual	0.02	0.08	0.20	0.02	0.07	0.25	0.02	0.10	0.16
	(0.13)	(0.28)	(0.40)	(0.14)	(0.26)	(0.43)	(0.13)	(0.30)	(0.37)
Experience in the current main occupation	9.89	5.51	4.32	10.66	6.02	3.62	9.20	4.50	4.76
	(9.10)	(7.24)	(6.17)	(9.27)	(7.78)	(5.56)	(8.88)	(5.93)	(6.50)
Training	0.15	0.14	0.07	0.17	0.14	0.08	0.14	0.12	0.06
	(0.36)	(0.34)	(0.25)	(0.38)	(0.35)	(0.27)	(0.34)	(0.33)	(0.23)
Union	0.72	0.29	0.02	0.73	0.31	0.01	0.70	0.25	0.02
	(0.45)	(0.45)	(0.13)	(0.44)	(0.46)	(0.12)	(0.46)	(0.43)	(0.14)

Note: Standard deviations are in parentheses.

Table 7.6 *Descriptive Statistics, Colored*

Variable	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Hourly wage	26.97	11.73	8.80	28.69	12.16	11.15	25.09	11.09	7.20
	(22.98)	(14.40)	(11.43)	(24.40)	(15.43)	(15.55)	(21.21)	(12.68)	(6.97)
Years of schooling	10.63	8.63	6.80	10.32	8.18	7.32	10.99	9.28	6.44
	(3.54)	(3.43)	(3.59)	(3.64)	(3.61)	(3.68)	(3.39)	(3.04)	(3.43)
Age	39.67	35.77	40.80	40.01	36.61	39.42	39.28	34.57	41.76
	(10.49)	(11.15)	(10.95)	(10.51)	(11.70)	(11.93)	(10.47)	(10.20)	(10.12)
Female	0.46	0.41	0.59						
	(0.50)	(0.49)	(0.49)						
Married	0.63	0.53	0.58	0.71	0.59	0.58	0.54	0.45	0.57
	(0.48)	(0.50)	(0.49)	(0.45)	(0.49)	(0.49)	(0.50)	(0.50)	(0.50)

Variable	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Permanent	0.83	0.72	0.51	0.85	0.73	0.33	0.81	0.69	0.61
	(0.37)	(0.45)	(0.50)	(0.36)	(0.44)	(0.47)	(0.39)	(0.46)	(0.49)
Fixed	0.07	0.05	0.03	0.06	0.05	0.05	0.09	0.05	0.02
	(0.26)	(0.22)	(0.16)	(0.25)	(0.22)	(0.21)	(0.28)	(0.21)	(0.12)
Temporary	0.05	0.06	0.11	0.04	0.07	0.13	0.06	0.04	0.11
	(0.22)	(0.23)	(0.32)	(0.20)	(0.26)	(0.33)	(0.24)	(0.20)	(0.31)
Casual	0.04	0.11	0.34	0.04	0.11	0.47	0.04	0.12	0.26
	(0.19)	(0.32)	(0.47)	(0.19)	(0.31)	(0.50)	(0.19)	(0.33)	(0.44)
Experience in the current main occupation	11.11	6.08	4.84	11.53	6.47	2.94	10.62	5.52	5.90
	(9.90)	(7.49)	(7.10)	(9.87)	(7.95)	(5.21)	(9.93)	(6.76)	(7.78)
Training	0.17	0.08	0.05	0.14	0.08	0.05	0.19	0.08	0.05
	(0.37)	(0.27)	(0.22)	(0.35)	(0.27)	(0.22)	(0.39)	(0.28)	(0.22)
Union	0.69	0.25	0.04	0.71	0.24	0.02	0.66	0.27	0.05
	(0.46)	(0.43)	(0.19)	(0.45)	(0.43)	(0.15)	(0.48)	(0.44)	(0.21)

Note: Standard deviations are in parentheses.

Table 7.7 *Descriptive Statistics, Asian*

Variable	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Hourly wage	42.89	29.91	20.57	46.14	32.92	23.17	35.70	23.70	12.31
	(35.84)	(33.19)	(25.66)	(40.50)	(37.41)	(27.97)	(21.59)	(20.87)	(14.28)
Years of schooling	13.01	11.71	10.10	12.85	11.59	10.16	13.29	11.94	9.93
	(2.30)	(2.17)	(2.76)	(2.44)	(2.15)	(2.86)	(2.03)	(2.20)	(2.50)
Age	38.38	38.09	38.60	39.85	38.87	38.57	35.79	36.54	38.71
	(9.49)	(10.96)	(10.81)	(9.77)	(11.31)	(10.76)	(8.49)	(10.07)	(11.38)
Female	0.36	0.33	0.24						
	(0.48)	(0.47)	(0.43)						
Married	0.72	0.71	0.62	0.78	0.74	0.66	0.63	0.66	0.50
	(0.45)	(0.45)	(0.49)	(0.42)	(0.44)	(0.48)	(0.49)	(0.48)	(0.52)

Variable	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Permanent	0.97	0.86	0.47	0.97	0.86	0.45	0.97	0.86	0.50
	(0.17)	(0.35)	(0.51)	(0.17)	(0.35)	(0.52)	(0.16)	(0.35)	(0.55)
Fixed	0.00	0.07	0.12	0.00	0.07	0.09	0.00	0.06	0.17
	(0.00)	(0.25)	(0.33)	(0.00)	(0.26)	(0.30)	(0.00)	(0.23)	(0.41)
Temporary	0.02	0.04	0.06	0.03	0.04	0.09	0.00	0.04	0.00
	(0.14)	(0.19)	(0.24)	(0.17)	(0.19)	(0.30)	(0.00)	(0.19)	(0.00)
Casual	0.01	0.04	0.29	0.00	0.03	0.27	0.03	0.05	0.33
	(0.10)	(0.19)	(0.47)	(0.00)	(0.18)	(0.47)	(0.16)	(0.22)	(0.52)
Experience in the current main occupation	9.54	7.32	6.06	10.51	7.57	3.64	7.84	6.89	10.50
	(8.15)	(7.74)	(7.77)	(8.44)	(8.06)	(4.84)	(7.43)	(7.17)	(10.46)
Training	0.13	0.13	0.07	0.14	0.13	0.07	0.11	0.13	0.07
	(0.33)	(0.33)	(0.26)	(0.35)	(0.33)	(0.26)	(0.31)	(0.34)	(0.27)
Union	0.66	0.37	0.06	0.73	0.37	0.10	0.53	0.36	0.00
	(0.48)	(0.48)	(0.25)	(0.45)	(0.48)	(0.32)	(0.51)	(0.48)	(0.00)

Note: Standard deviations are in parentheses.

Table 7.8 *Number of Workers by Sector and Industry*

Main industry	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Agriculture, hunting, forestry and fishing	54	2,551	1,053	35	1,769	576	19	782	477
Mining and quarrying	21	872	5	20	794	4	1	76	1
Manufacturing	92	2,404	458	72	1,625	209	19	779	249
Electricity, gas and water supply	82	72	7	64	59	7	18	13	0
Construction	133	1,153	707	67	1,010	634	66	142	73
Wholesale and retail trade	92	3,599	1,882	50	1,838	705	42	1,760	1,176
Transport, storage and communication	201	543	199	154	450	174	47	93	24
Financial intermediation, insurance, real estate and business services	166	1,486	108	94	846	67	71	639	41
Community, social and personal services	3,642	1,107	298	1,638	413	117	2,002	693	181
Private households	12	108	2,350	4	24	487	8	84	1,862
Exterior organisations and foreign government	3	5	0	1	4	0	2	1	0
Unspecified	1	12	9	1	6	6	0	6	3
Total	4,499	13,912	7,076	2,200	8,838	2,986	2,295	5,068	4,087

Table 7.9 *Number of Workers by Sector and Occupation*

Main occupation	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Legislators, senior officials and managers	218	1,016	161	151	706	105	67	310	56
Professionals	784	482	37	297	265	18	484	217	19
Technical and associate professionals	1,128	854	154	376	459	78	752	395	76
Clerks	565	1,505	47	204	435	16	361	1,067	31
Service workers and shop and market sales workers	672	1,687	636	415	922	204	256	764	432
Skilled agricultural and fishery workers	23	323	757	18	253	360	5	70	397
Craft and related trades workers	232	2,000	1,123	209	1,696	837	23	304	286
Plant and machine operators and assemblers	183	1,951	206	161	1,633	183	22	317	22
Elementary Occupation	683	3,994	1,976	366	2,453	1,050	317	1,540	925
Domestic workers	9	87	1,969	1	8	129	8	79	1,839
Unspecified	2	13	10	2	8	6	0	5	4
Total	4,499	13,912	7,076	2,200	8,838	2,986	2,295	5,068	4,087

Figure 7.1. Distribution of Wage by Sector

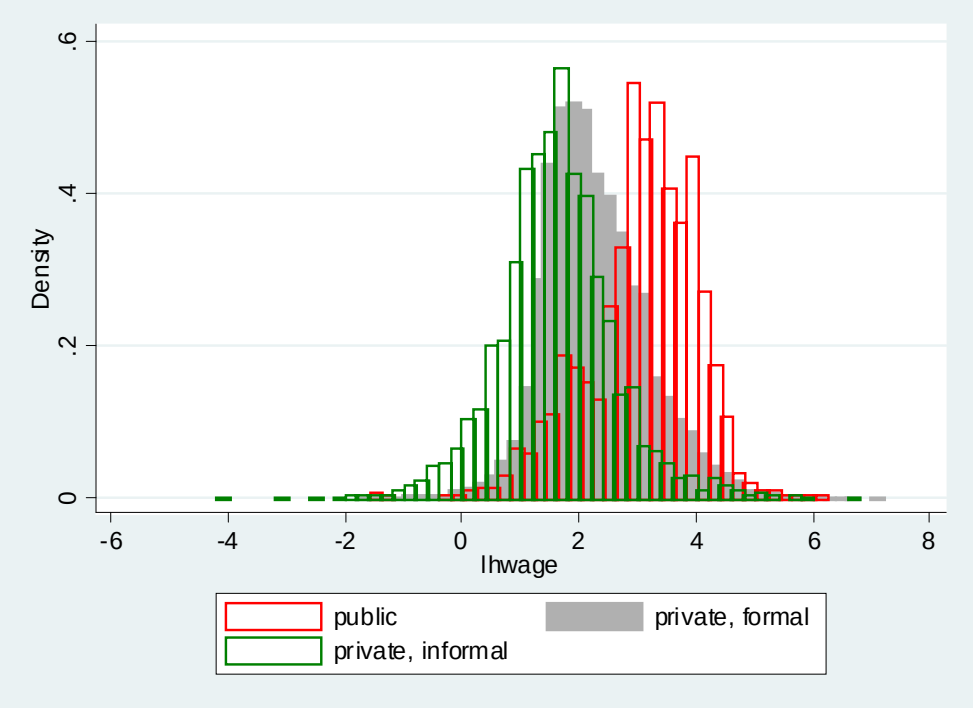


Figure 7.2. Distribution of Years of Schooling by Sector

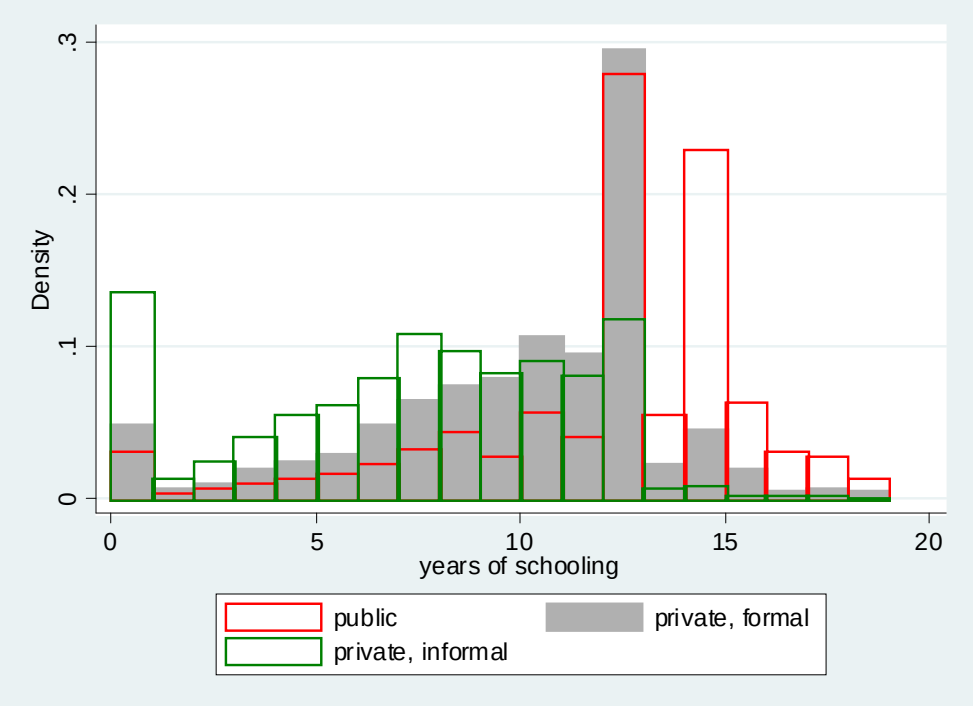
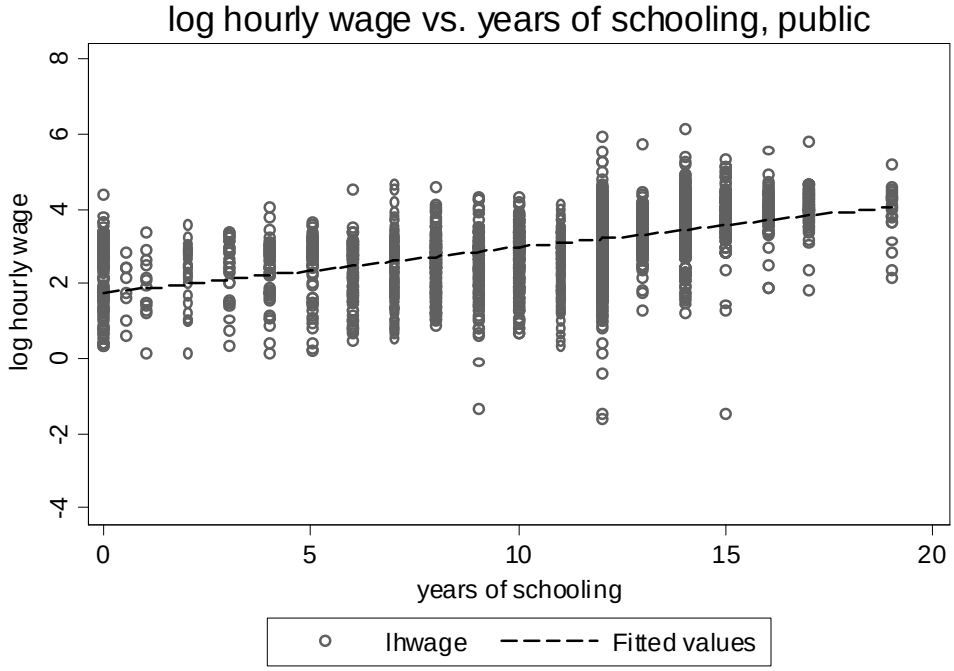


Figure 7.3. Years of Schooling vs. Log Hourly Wage by Sector



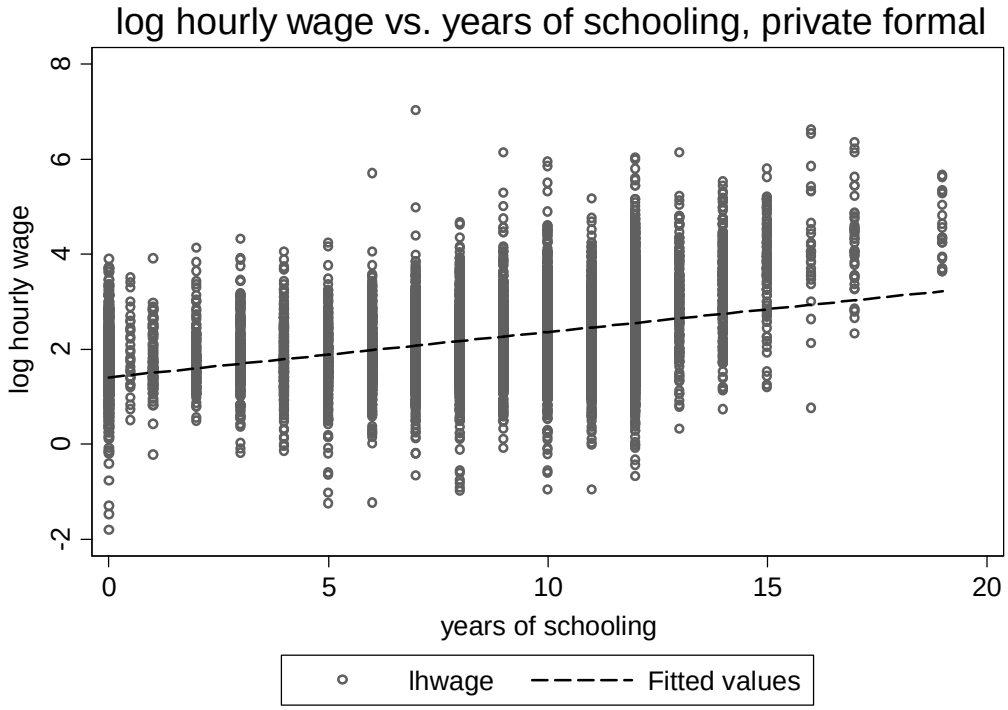


Figure 7.3 (continued)

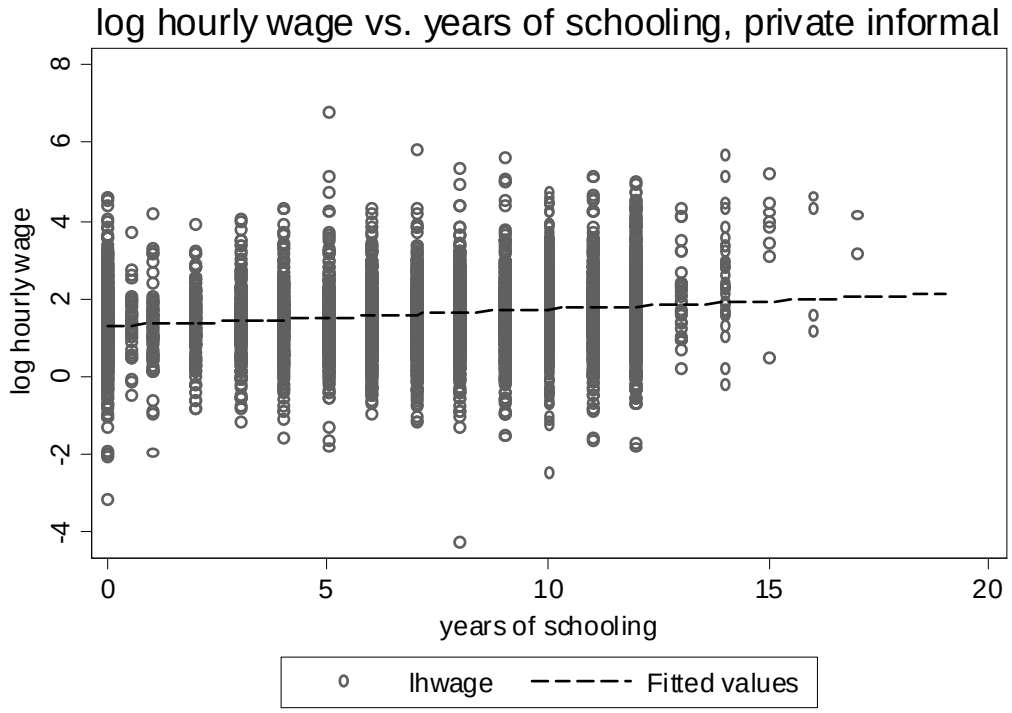
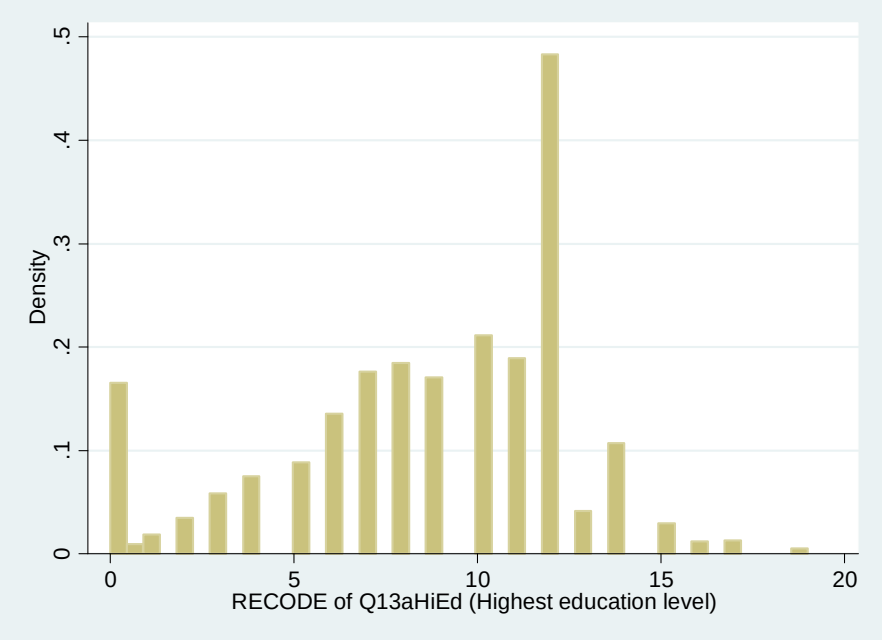
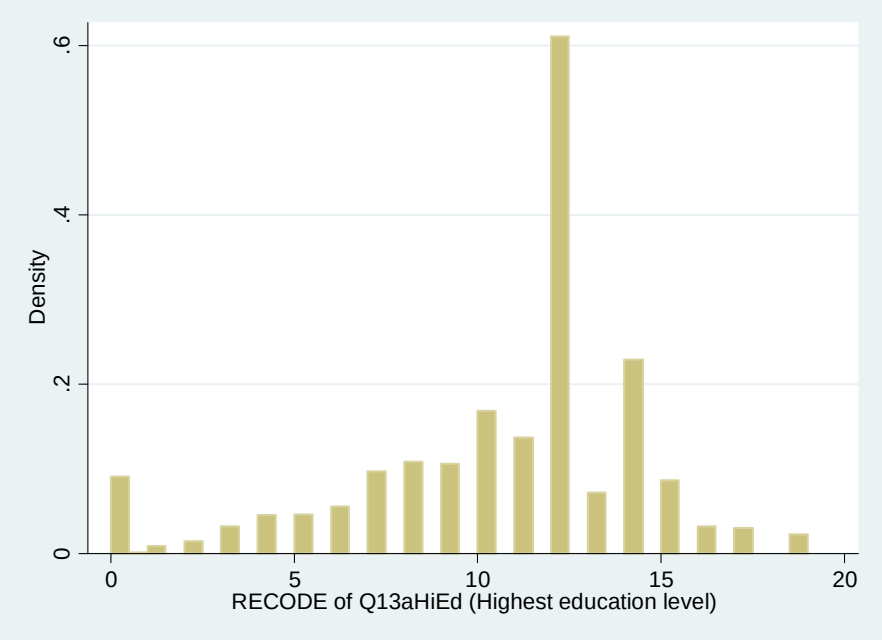


Figure 7.4. Distribution of Years of Schooling



(The Sample of Workers Who Reported Earnings)



(The Sample of Workers Who Did Not Report Earnings)

Note: Workers here means those who reported whether they work at either the formal or informal sector.

8. RESULTS OF THE ESTIMATION OF MARGINAL RETURN TO SCHOOLING

This section presents the results of the estimation of the marginal returns to schooling. The following analysis was conducted: 1) ordinary least squares (OLS); 2) two-stage least squares (2SLS) to control for endogeneity of schooling; 3) two-step selection correction models; and 4) the combination of 2SLS and two-step selection correction model controlling for both endogeneity of education and sample selection bias from sector choice and labor force participation. The regressions were conducted for the three sectors identified before: public, formal private and informal private sectors, and was disaggregated for males and females, and for different races. Although the results are presented for all four models examined, for comparability and robustness purposes, the fourth model is the most sophisticated one because it corrects for both the endogeneity of education and sector sample selection bias.

8.1. OLS Results

The simplest method used was ordinary least squares (OLS), carried out by sector, gender, and race. Table 8.1 shows the coefficient of years of schooling, its robust standard error, and the number of observations in each OLS regression.[1] The coefficients of years of schooling are all significant at the 1% or 5% level, except for a few cases where sample sizes were small. (the results for the informal private sector for male and female Asians are not reported because of the very small sample size).

Columns (1), (2), and (3) display the results for the total sample. The marginal return to schooling in the public sector is 13.7%, the highest among the three sectors. The marginal return

[1]OLS results for married people are also shown in Table 10.2 to compare with 2SLS results in the next section.

to schooling in the formal private sector is 9.6%, which is approximately 4% points lower than the one in the public sector. The marginal return to schooling in the informal private is 4.4%, a third of the one in the public sector and less than half of the one in the formal private sector. These results thus show positive returns to schooling in the informal private sector, but much lower than the ones in other sectors.

The results are similar for males and females. The marginal returns to schooling in the public, formal private, and informal private sectors are, respectively, 11.4%, 9.5%, and 5.3% for males and 16%, 10%, and 3.9% for females. Therefore, the difference in the marginal return to schooling between the public and formal private sectors is 4.1% overall, 1.9% for males, and 6% for females; the difference between the public and informal private sectors is 9.3% overall, 6.2% for male, and 12.1% for female; the difference between the formal and informal private sectors is 5.2% for total, 4.2% for males, and 6.1% for females. Therefore, the gaps in rates of return to education between sectors are greater for females.

The patterns in marginal returns to schooling among different sectors by race are very similar to those for the total, male and female, samples, although there are some differences for specific population groups. Also, we cannot make good inferences about male and female Whites and male and female Asians because of the very small sample sizes available for these groups. The difference in marginal returns to schooling between the public and formal private sectors (the marginal return to schooling in the public sector minus the one in the formal private sector) is generally positive and ranges from -1.7% to 17.2%. The differential in marginal returns to schooling between the public and informal private sectors varies from 5.8% to 12.3% in subgroups by race and gender. The difference in marginal return to schooling between the formal

and informal private sectors is between 3.4% to 5.7% in subgroups by race and gender. [2]

The returns to schooling in the public sector are the highest among Whites, followed by Asians. The marginal return to schooling in the public sector is 18.5% for Whites, 13.6% for Blacks, 13.8% for Coloreds, and 17.1% for Asians. The returns to schooling in the formal private sector are also highest for Whites, followed by Asians,.and then Blacks and Colored. The marginal returns to schooling in the formal private sector are 16.4% for Whites, 8.8% for Blacks, 10.7% for Coloreds, and 14.7% for Asians. The marginal returns to schooling in the informal private sector are 4.2% for Blacks and 5.7% for Colored. The marginal returns to schooling for Whites and Asians are 9.4% and 4%, but these are statistically insignificant partly because the sample sizes for White and Asian workers in the informal sector are very small.

When disaggregated by gender, the results for males and females are very similar to the overall results just reported. There are, however, some differences. The marginal return to schooling in the public sector is higher for males than females for Whites, while it is lower for Blacks, Coloreds, and Asians.

8.2. Results of the Two-Stage Least Squares Controlling for Endogeneity of Education

Next, the results of the two-stage least squares (2SLS) analysis are reported, conducted to control for the endogeneity of education. First, tests to check the validity of the instrument are shown. Then, the 2SLS estimation results are presented.

8.2.1. Tests for the instrument

Table 8.3 shows the statistics to test the endogeneity of years of schooling for the overall sample. Specifically, it presents the results of the Wu-Hausman F test, and the test of excluded

[2]White and Asian samples were excluded from the analysis of the difference in marginal return to schooling between the public and informal private sectors, and between the formal and informal private sectors here.

instrument (or weak identification test).

The Wu-Hausman test examines the endogeneity of years of schooling. It checks whether the coefficient of the residual is statistically different from zero. It also examines the null hypothesis that that education is exogenous (i.e., there is no statistically significant difference between the OLS and IV estimates). The ordinary least squares estimator fails if there is correlation between an explanatory variable and the error term. The instrumental variables estimator can be used when the ordinary least squares estimator has this problem. Hausman (1978) suggested comparing the OLS and 2SLS estimators of the coefficient of a variable of interest as a formal test of endogeneity: if the y_2 is uncorrelated with u_1, the OLS and 2SLS estimators should differ only by sampling error.

The test of excluded instruments is the test of weak instrument. It calculates the F-statistics for the null hypothesis that the coefficient estimates of the instrumental variable in the first-stage regression are not different from zero. This test is important because when the partial correlation between the instruments and the endogenous variable is weak, instrumental-variables regression is biased in the direction of the OLS estimator (e.g., Staiger & Stock, 1997). Staiger and Stock (1997) recommend that *F*-statistics (or the corresponding *p*-values) from the first-stage regression be routinely reported in the studies using instrumental variables. The *F*-statistic tests the hypothesis that the instruments should be excluded from the first-stage regressions (i.e., the relevance of the instruments). The logic is that when the *F*-statistic is small (or the *p*-value is large), the instrumental variable estimations and the corresponding standard errors are unreliable (Bound, Jaeger, & Baker, 1995; Staiger & Stock, 1997; Stock & Yogo, 2004).

The weak identification test is similar to the test of excluded instrument, but its use is

recommended when one estimates the General Moment Method (GMM). For example, if one calculates the robust standard error, one needs to use the weak identification test instead of the test of excluded instrument. The weak identification test uses the Kleibergen-Paap rk Wald F-statistic and evaluates its statistical significance based on the Stock-Yogo weak ID test critical values. In this analysis, the Kleibergen-Paap rk Wald F-statistic is the same as the F-statistics for the test of excluded instrument.

It should be noted that the model has one instrument for one endogenous variable, and therefore the model is just identified. This means one cannot conduct the overidentification test because overidentification tests can be conducted only when there are more instruments than endogenous variables (Wooldridge, 2002b).

As indicated in Table 8.3, the Wu-Hausman test shows that there are statistically significant differences between OLS and 2SLS estimates and that years of schooling is the endogenous variable for the public, formal private, and informal private sectors, respectively, for the samples including gender and all races, as well as the male sample and female sample at less than the 2.5% level. Also, according to Table 8.4 and Table 8.5, statistically significant differences between OLS and 2SLS estimates exist for most of the estimation by sector, race, and gender at the 5% level, with some minor exceptions due to small sample sizes.

The tests for excluded instrument show strong correlations between the instrument (spouse's years of schooling) and years of schooling for the total sample, male sample, and female sample for all races and most of the sub-samples by race and gender. The F-statistics range from 7.92 to 939.69, most of which are well above the critical value, 10, which was suggested by Stock and Watson (2003, p348-372) as being a critical value. Spouse's years of schooling appears to be weakly correlated with years of schooling only for Asian males and

females working in the public sector. The F-statistics are 7.92 for Asian males working in the public sector and 8.35 for Asian females working in the public sector. Also the F-statistics are above the critical values for weak instruments, as recommended by Stock, Wright, and Yogo (2002). Therefore, the instrument used is strong enough to make an inference for the finite sample, except for some specific sub-samples by race and gender.

8.2.2. Two-stage least square estimations

Next, the results of the estimation using 2SLS are reviewed. Table 8.5 shows the coefficients on years of schooling, with robust standard errors, and the number of observations in each 2SLS regression.. As explained in the methods section, the model uses the spouse's years of schooling as the instrumental variable. Therefore, the sample covered in the model is only the group of married people.

For comparison, the results of OLS estimation by sector, gender, and race for married people are shown in the Table 8.2. The OLS results for married people are not much different from those for both single and married. Again, the coefficients for the years of schooling variable are mostly significant at the 1% level, with just a few exceptions for sub-groups with small observations.

The 2SLS coefficients for the rate of return to education are several percentage points higher than the OLS estimates across different sectors, gender, and race. This suggests that in this South African sample, there is a downward bias in the return to schooling because of unobserved individual characteristics such as ability, which is the opposite of the theory developed by Card (1995, 1999), as presented in Chapter 5. Also, the marginal return to schooling in the formal private sector is now greater than that in the public sector, both for total sample and many sub-samples. The marginal return to schooling in the informal private sector is the lowest across

gender and race, and this result remains the same as in the OLS estimations.

These are the specific results obtained for the total sample, which includes all races. The marginal return to schooling in the public sector is now 17.0%, while it is 12.3% for the OLS estimation for married people and 13.7% for the OLS estimation for both single and married people. The 2SLS estimate is 4.7 percentage points higher than the OLS estimates for the married sample, and 3.3 percentage points higher than the OLS estimates for the total sample. The marginal return to schooling in the formal private sector is 19.5% for the 2SLS, compared to 10.2% for the OLS for married people and 9.6% in the OLS for the total sample. Therefore, the 2SLS estimate is 9.3% higher than the OLS estimates for married and 9.9% higher than the OLS estimates for total. The rate of return in the informal private sector for 2SLS is 12.3%, which is also much larger than OLS estimations (5.2% for the sample of married people and 4.4% for the total sample). The 2SLS estimate is 7.1% higher than the OLS estimates for married and 7.9% higher than the OLS estimates for total. The downward bias in the OLS estimate is the largest in the formal private sector, followed by the informal private sector. Also, the marginal returns increased for the sub-samples by gender and race as well. The marginal return to schooling in the formal private sector becomes greater than the one in the public sector.

8.3. Two-Step Selection Correction Model

Next, the results of the two-step selection correction models are discussed. As reviewed in Chapter 7, several selection correction methods are available. The general finding is that the marginal return to schooling in the formal private sector is higher than the one in the informal private sector, although the difference tends to be smaller compared to 2SLS and OLS in some of the two-step selection correction models (DMF variant 1, DMF variant 1, and Dahl 2). This is true across race groups, except for White males and females and Asians, for which there are

insufficient observations. Also, the marginal return to schooling in the public sector becomes smaller after controlling for sample selection bias, compared to the OLS and 2SLS. In the choice equation here, the level of education dummies are used instead of the years of schooling continuous variable. This is because multicollinearities seem to occur between the selection correction terms and 2SLS, which is the next model. To compare with the next model, the two-step selection correction model uses the level of education dummies.

8.3.1. Tests for the validity of the exclusion restriction

In order to meet the exclusion restriction, spouse formal private sector, spouse informal private sector, household formal private sector, and household informal private sector dummies were added in the choice model.[3] The assumption is that the types of sector of employment for spouse and household members would affect the choice of the sector of workers, but would not directly affect the earnings of workers. There is no established method to test the exclusion restriction in the two-step selection correction model, but McEwan (2001) used informal specification tests. McEwan (2001) checked whether the coefficients of (possible) instruments in the choice model were large enough and statistically significant; and added the instruments in the second stage equation to see if they had effects on the dependent variable.

Table 8.6 shows the marginal effects and standard errors of spouse formal private sector, spouse informal private sector, household formal private sector, and household informal private sector dummies. These indicate that the sector of employment for spouse and household members do affect the sectoral choice of workers. The correlation between spouse informal

[3]Putting in six sector dummies, including spouse public sector and household public sector dummies, was also considered. However, they caused a conformability error. Therefore, spouse public sector and household public sector dummies were removed.

private, household formal private, household informal private, and sector choice was statistically significant across sector, with the exception of household formal private and household informal private for the formal private sector choice. Also, the joint effect of these dummies was statistically significant. Therefore, one can assume that the sector of employment for spouse and household members affect the sectoral choice of workers.

Next, wage equations were estimated, including spouse and household dummies as additional explanatory variables. The significant correlations between these dummies and earnings suggest that the exclusion restriction is not valid. Table 8.7 and Table 8.8 show the coefficients of spouse and household sector dummies and the joint significance of their effects in wage equations. Most of the coefficients are not statistically significant across sector, race, and gender. Also, the joint significance tests indicate that one cannot reject the hypothesis that the coefficients of spouse and household sector dummies were jointly equal to 0 for most of the groups. The results show that correcting for selection bias could be less appropriate for total in the public sector, total in the formal private sector, males in the formal private sector, females in the public sector, total Blacks in the public and formal private sectors, male Blacks in the formal private sector, female Blacks in the informal private sector, total Coloreds in the public sector, and Colored females in the public sector.

8.3.2. Results

Table 8.9 shows the estimation results of the two-step selection model using different selection terms for total, male, and female. Also, Figure 8.1 shows the coefficients of marginal return to schooling in different sectors in different models. The two-step selection correction models estimated here only include married people in order to compare with the 2SLS, which has only married people in the sample. The coefficients in the models were compared mainly

with the OLS results.

After controlling for sample selection bias, the magnitudes of coefficients became much smaller in the public sector across models, compared to the OLS estimations. The estimated marginal returns to schooling were unchanged in the formal private sector for Lee, DMF, and Dahl selection correction models, while they became much smaller for DMF variant 1, DMF variant 2, and Dahl 2, compared to the OLS results. The coefficients for the informal sector did not change or slightly changed for Lee, DMF, and Dahl 1. However, they decreased by 0.2 to 1.7% points.

After correcting for sector selection bias, the difference in the marginal return to schooling between the public and formal private sectors became much larger for the Lee, DMF, and Dahl 1 model, while it became smaller for the DMF variant 1, DMF variant 2, and Dahl 2, compared to the OLS. The gap ranged from 3.6% to 5.2% for the former models, and from 0.8% to 2.7%. The difference in the marginal return to schooling between the public and informal private sectors became much smaller across models, and the gap ranged from 0.3% to 2.6% across models. The difference between formal and informal private remained unchanged for Lee, DMF, and Dahl 1 models, with the gap ranging from 4.3% to 6.3%. It became smaller for DMF variant 1, DMF variant 2, and Dahl2; the difference was between 2% to 4.7% for the three models. In any event, the difference in return to schooling between the formal and informal private sectors still existed after controlling for the sample selection bias.

The marginal return to schooling in the formal sector remains the highest among all sectors in the two-step selection models, except for the female sample where the marginal return to schooling in the public sector was the highest. To sum, the effect of the endogeneity of schooling was smaller than the effect of sector sample selection bias in the public sector. In

contrast, the effect of the endogeneity of schooling was much stronger than the effect of sector sample selection bias in the formal private sector, especially when we used Lee, DMF, and Dahl 1 selection correction. The endogeneity of schooling affected the return to schooling to a greater degree in the informal sector, while the sample selection bias did not affect it very much.

Tables 8.10-8.15 display the marginal return to schooling by race and gender estimated by the Lee, DMF, DMF variant 1, DMF variant 2, Dahl 1, and Dahl 2 two-step selection correction methods. The findings for Blacks were the same as the sample of all races. The changes between the OLS/2SLS and the two-step selection correction for Coloreds were also similar to the estimations for the sample of all races, except that the marginal return to schooling in the informal sector became lower in DMF variant 1, DMF variant 2, and Dahl 2 models, compared to the OLS. In contrast, the results for Whites and Asians were very different. For example, the marginal return to schooling in the public sector was not much different from the OLS and 2SLS for the total White sample across models, while it usually had the opposite trend for the total sample and other groups. For the White male sample, the marginal return to schooling in the public sector was also similar to the OLS and 2SLS, but was much lower in the model with the Dahl 2 selection correction. The marginal returns to schooling in the public sector in the two-step selection correction models were similar to OLS, but much lower than the 2SLS for White females. Although the trends of change for total and male Asians were similar to the sample of all races, female Asians had different trends. The marginal return to schooling in the public sector in the two-step selection correction for female Asians was similar to that in the OLS and 2SLS, except for the DMF variant 1 estimation, which was much lower than the OLS and 2SLS, and the Dahl 1, Dahl 2, and DMF variant 2 estimations, which were much higher than the OLS and 2SLS.

As we have seen, the rates of return to schooling in the formal sector were usually very different between the Lee/DMF/Dahl 1 and DMF variant 1/DMF variant 2/Dahl 2. Based on discussions by Bourguignon et al. (2007) and the observed multicollinearity between some selection correction terms and the two-stage least square, the study used the DMF variant 1 selection correction as the main selection correction term. Although the next section shows the results of 2SLS with all selection corrections, only the DMF variant 1 was used for the quantile regression, the piecewise linear spline function, and a robustness check.

8.4. Results of the 2SLS with Selection Correction Estimation

Table 8.16 displays the results of the 2SLS/two-step selection estimation model. This model controls for both endogeneity of education and sector sample selection bias. Therefore, this model provides the most comprehensive estimate of the marginal rate of return to schooling in different sectors obtained in this dissertation. As pointed out earlier, several sample selection correction methods are available. For this analysis, different models with different selection corrections were estimated. It should be noted that the level of education dummies were used instead of the years of schooling continuous variable in the choice equation because multicollinearities seem to occur between the selection correction terms and 2SLS. Figure 8.6 shows the results of different models including the OLS and 2SLS.

The estimated marginal returns to schooling may vary from method to method used. As indicated by Bourguignon et al. (2007), the Lee selection correction works for only small samples, and the Dubin and McFadden variant 1, Dubin and McFadden variant 2, and Dahl 2 work better. However, the Dahl 2 correction method seems to lead to multicollinearity in smaller samples such as sub-samples by race and gender because the selection correction consists of several terms when the estimation of wage equation is conducted with 2SLS and selection

correction. Bourguignon et al. (2007) also mentions that the DMF variant 2 is less robust than the DMF and DMF variant 1.

Table 8.16 and Figure 8.6 indicate that, overall, the marginal returns to schooling does not change much in the public sector for the total sample and for males across the various models, when compared to the 2SLS, while it became even higher than 2SLS estimations (i.e., also higher than OLS estimations) for females across all models. The marginal return to schooling in the formal private sector does not change for the total sample, for the male and female samples in the 2SLS models with the Lee, DMF and Dah1 selection corrections. However, it is much higher for the total, male and female in the 2SLS models with the DMF variant 1, DMF variant 2 and Dahl 2 selection correction. The marginal return to schooling in the informal private sector does not change significantly for the total, male and female samples in the 2SLS with the Lee, DMF, and Dahl 1; it did became higher for the three samples in the DMF variant 1, DMF variant 2 and Dahl 2.

Tables 8.17-8.22 display the marginal returns to schooling estimated by the 2SLS with different selection correction methods by race and gender. Also, Figures 8.7-10 show the estimated marginal return to schooling in different models including the OLS and 2SLS for different race groups and by gender. The changes in the estimates of the return to schooling between the OLS and 2SLS, and the 2SLS with selection correction for Blacks are similar to the changes observed for samples of all races except that the marginal return to schooling in the public sector is higher in the 2SLS with any selection correction compared to 2SLS for total Black and that the increase in the marginal return to schooling in the public sector is much sharper across different models with different selection correction.

The marginal returns to schooling in the public sector estimated by the 2SLS with Lee,

DMF, and Dahl 1 selection correction becomes lower than the one estimated by the 2SLS and slightly higher than the OLS estimates for total Colored; they get a few percentage point higher than the 2SLS for male Colored; and they became negative but statistically insignificant for Colored female. The marginal returns in the public sector estimated by the 2SLS with DMF variant 1, DMF variant 2, and Dahl 2 selection corrections is lower than the 2SLS estimate and similar to the OLS estimate for total Colored, while they are even higher than the 2SLS estimate for male Colored. They are again negative and statistically insignificant for female Colored. The estimations in other sectors for Colored are similar to the ones for the overall sample.

The estimated marginal returns to schooling in the public sector in the 2SLS with different selection corrections are much higher than the 2SLS estimation for total Whites. Also the marginal returns to schooling estimated by the 2SLS with DMF variant 1, DMF variant 2 and Dahl 2 selection corrections are lower than the 2SLS and OLS estimates for White male, especially the one estimated by the 2SLS/Dahl 2 is much lower than other estimates for White male. The marginal returns to schooling in the public and formal private sector in the 2SLS/DMF variant 1, 2SLS/DMF variant 2, and 2SLS/Dahl 2 are much greater than the ones estimated by the 2SLS and OLS.

Asians have very different rates of return when compared to other groups. First of all, the marginal returns to schooling in the public sector for Asians are all statistically insignificant when estimated by the 2SLS with selection corrections. Although they are statistically insignificant, the changes in the magnitude between models are explained here. The marginal returns to schooling in the public sector estimated by the 2SLS with selection corrections are similar to the OLS and 2SLS estimates for Asian total; they are lower than the OLS and 2SLS estimates for Asian male. The marginal returns to schooling the public sector estimated by the

2SLS with the Lee and DMF are lower than the OLS and 2SLS estimates for Asian female, while the ones estimated by the 2SLS with DMF variant 1, DMF variant 2, Dahl 1 and Dahl 2 is much greater than the OLS and 2SLS estimates for Asian female. The marginal returns to schooling in the formal sector for Asians estimated by the 2SLS with Lee, DMF and Dahl 1 are statistically significant while the ones estimated by the 2SLS with DMF variant 1, DMF variant 2 and Dahl 2 are not statistically significant.

The marginal return to schooling is higher in the formal private sector than the one in the public sector for most of subgroups as we also observe in the 2SLS estimations except for the one in for female estimated by the 2SLS with Lee, DMF, DMF variant, DMF variant 2 and Dahl selection correction, White total, White female, Black total estimated by the 2SLS with Lee, DMF and Dahl 1, Black female and Asian female estimated by the 2SLS with all selection corrections. Also the marginal return to schooling in the informal private sector is still lower than the one in the public sector even both endogeneity of education and sample selection bias are taken into account. Furthermore the return to schooling in the informal sector is the lowest across different groups except for most of the 2SLS/selection correction models for Colored total, male and female.

Table 8.19 displays the 2SLS estimates with the DMF variant 1 selection correction model in more detail. The marginal return to schooling in the public sector for the total samplel is now 18.1%; one percentage point higher than the 2SLS estimate, and 4.4 percentage points higher than the OLS estimate. The marginal return to schooling in the formal private sector is 25.1%; 5.6% points higher than the 2SLS, and 15.5% higher than the OLS. The marginal return to schooling is now 16.8%, which is 4.5% points higher than the 2SLS, and 12.4% points higher than the OLS. The trend is similar across gender and race. However, the positive impact of

selection correction is much greater for Whites in the public and formal private sector, and Black females in the public sector. The marginal return to schooling in the public sector for the overall White sample is now 31.3%, compared to 22.4% in the 2SLS and 18.5% in the OLS. The one in the formal private sector for Whites is 49.2%, compared to 26.0% in the 2SLS and 16.4% in the OLS. The marginal return to schooling in the public sector for Black females is 40.4%, 16.3% points greater than the one in the 2SLS, and 24.6% points greater than the one in the OLS.

The difference in the marginal return to schooling between the public and formal private sector is -7.0% for total, -9.7% for males, and 1.8% for females; the gap is greater than the 2SLS estimate for the total and male samples. The difference in the marginal return to schooling between the public and informal private sector is 1.3% for total, 0.2% for male, and 8.4% for female, which is smaller than the 2SLS estimates. The differential in the marginal return to schooling between the formal and informal private sector is now 8.3% for total, 9.9% for male, and 6.6% for female in the sample including all races. Looking the results by race, the difference in returns to schooling between the formal and informal private sectors is 7.3% for Black total, 11.1% for Black male, 3.4% for Black female, 13.2% for Colored total, 10.9% for Colored male, and 16.1% for Colored female. The difference is especially large for Colored female.

The findings in the 2SLS with the selection correction model above confirm the substantial difference in the rate of return to schooling between the formal and informal private sectors, as well as the differences between the public and formal private sector. The difference in return to schooling between the public and informal private sector also exist but it is much smaller than other comparisons of sectors.

8.5. Robustness Check

We have already found that the lower return to schooling in the informal sector compared to the public and formal sectors was robust across gender and race in the models employed above. This section pursues a further robustness check on the relative size of returns to schooling by sector. First, the 2SLS estimation with the DMF variant 1 selection correction was conducted using different definitions of the informal sector. Second, the researcher conducted the 2SLS estimations with DMF variant 1 selection correction model with district dummies, occupation dummies, and industry dummies.

First, the definition of the informal sector was changed based on the definition of the informal employment by Statistics South Africa (2010), introduced in Chapter 2. This newly-used definition focuses more on informal employment, not on working in the informal sector. According to Statistics South Africa (2010), informal workers (not workers in the informal sector) are workers in the informal sector, employees in the formal sector, and individuals working in private households who lack entitlements to basic benefits such as pension or medical aid and do not have a written contract of employment. We used three variations of the dependent variable based on this definition. The first definition is “working in the informal private sector and/or working without a written contract in the private sector.” Because the nature of the public sector is different from that of the private sector, workers without written contracts in the public sector were not included in the category above. Instead, a new category, the public informal employment, was introduced. The second definition of informal employment in the private sector is “working in the informal private sector, without any of medical aid/health insurance and pension/retirement fund” in the private sector. The third definition of informal employment in the private sector is “working without any of the following:

written contract, medical aid/health insurance, and retirement fund."

Table 8.23 presents the results of the 2SLS estimation with DMF variant 1 selection correction using the definition of the informal employment as those who work in the informal sector and/or without written contract. Using this definition, the difference in returns to schooling between the formal private and informal private sectors becomes smaller but is still significant, except for Black females where the difference is only 0.8%.

Table 8.24 presents the results of the 2SLS estimation with DMF variant 1 selection correction using the definition of informal employment as those who are working in the informal sector and/or without any written contract, medical aid/health insurance or pension/retirement fund. The difference in returns to schooling between the formal private and informal private sectors is still substantial.

Further analysis was carried out using yet another definition of informal employment. These results are reported in Table 8.25. The definition of the informal employment is now those who lack any of the following: written contract, medical aid/health insurance, and pension/retirement fund. Again the difference in returns to schooling between the formal and informal sector is significant.

Table 8.26 displays the coefficient of the schooling variable in the 2SLS model with the DMF variant 1 with additional explanatory variables. The patterns in the return to schooling were similar to those reported earlier even after controlling for district, industry, and occupation. There is still a significant difference in return to schooling between the formal private and informal sectors. The findings are robust across race and gender (See Tables 8.26-30).

In additional analysis carried out, self-employed, head of company dummies were included in the model as well, but were dropped in all samples because of collinearity.

Self-employed and head of company are highly correlated with other variables (such as trade union participation and permanent status) already in the equation.

8.6. Results of the Wage Decomposition Analysis

The results in the previous sections showed the gap between returns to schooling in the informal private sector and the formal private sector and/or the public sector. Even controlling for the endogeneity of education and sample selection into sectors, we find a significant gap between the informal and formal sectors across gender and race. This seems to validate the dual market theory stated earlier: there appear to be two labor markets with different structures and mechanisms, including the compensation they provide to the educational attainment of workers. However, which variables explain the most the wage gaps between the formal and informal sectors? Is it possible that the wage gap is not due to differences in the characteristics of workers employed in each sector and is mostly responsive to the rewards that are offered to education or other characteristics in the two sectors?

The Blinder-Oaxaca decomposition was employed with the 2SLS and Heckman selection correction models used to control for the endogeneity of education and sample selection into sectors. To summarize: the results show that the lower coefficients on the education variable in the informal sector played a more important role in explaining wage differentials between the formal and informal sector than the differences in average schooling endowments between workers in the two sectors. On the other hand, differences in average schooling tended to contribute much more to wage differentials between the public and private sectors.

Table 8.30 shows the results of the Blinder-Oaxaca wage decomposition, indicating the contribution of differences in worker's average educational endowment relative to the contribution of differences in the returns to education in explaining the earnings differentials in

the 2SLS with the Heckman selection correction for the sample of all races. The role of differences in the workers' average education endowment was greater than the role of differences in the return to schooling, when comparing workers in the public and formal private sectors and workers in the public and informal private sectors. In contrast, the differences in the returns to schooling had a greater impact on the wage differentials between the formal and informal private sectors, compared to differences in workers' average education endowment. This result also generally held for the male and female samples. except for the wage differentials between the public and formal private sectors for male workers. The differences in returns to schooling affected the wage differentials between the formal and informal sectors than the differences in average education endowment.

Tables 8.31 presents the contribution of differences in workers' average educational endowments relative to the contribution of differences in the returns to education to the earnings differentials in the 2SLS with the Heckman selection correction derived from of the Blinder-Oaxaca wage decomposition for the samples of different races. The estimation of wage differentials between the public and informal private sectors and between the formal private and informal private sectors are not available for Whites and Asians because of insufficient observations. Despite some exceptions, the trends for Blacks and Coloreds were similar, especially for wage differentials between the formal private and informal private sectors: the role of differences in the return to schooling was more important for wage differentials between the formal private and informal private sectors, compared to the role of differences in average education endowments.

8.7. Heterogeneity of Returns to Schooling

8.7.1. Results of the estimation of quantile regressions

This section presents the results obtained from the quantile regressions model. Table 8.32 and Figure 8.1 show the marginal returns to schooling across different points of the (conditional) wage distribution in different sectors, estimated by ordinary quantile regression, instrumental variable quantile regression, and instrumental variable quantile regression with selection correction of sector choice. To summarize: these results show that the conditional distributions of earnings given years of schooling were significantly different between the formal private and informal private sectors and do not often overlap in the model controlling both endogeneity of schooling and sample selection bias. The conditional distributions of earnings given years of schooling were similar between the public and informal private sector for the total sample and for males, while they are different for females across racial groups. The conditional distribution of earnings given years of schooling was very different between the public and formal private sectors across gender and race.

The results of the ordinary quantile regression equations are presented in Columns (1), (2), and (3) in Table 8.32 and the first figure in Figure 8.1. Returns to schooling in the public sector were the highest at the 10^{th} quantile, and then became lower as the quantile became higher. The marginal return to schooling at the 10^{th} quantile was 15.8% while the one at the 90^{th} conditional quantile was 12.8%. Returns to schooling in the formal private sector was the highest at the highest point of the wage distribution, and then became lower as the quantile decreased. The marginal return to schooling at the 10% quantile was 6.6% while the one at the 90% quantile was 11.6%. The marginal return to schooling in the informal private sector was the lowest at the 25^{th} conditional quantile (3.6%), while it was the highest at the 10^{th} and 90^{th} conditional quantiles

(5.2%). The difference between the highest and lowest marginal return to schooling across quantiles in the formal private was the largest. The difference between the highest and lowest marginal return to schooling across quantiles in the informal private sector was the lowest: the distribution was much flatter than the distributions of other sectors. The differences in rates of return to education among the extreme quintiles in the public, formal private, and informal private sectors were 3 percentage points, 4.9 percentage points, and 1.5% percentage points, respectively.

When the quantile regression equation was run with an instrumental variable to control for the endogeneity of education, the marginal return to schooling became higher in general (see columns (4), (5), and (6) in Table 8.32 and the first and second figures in Figure 8.1), as was seen in the two-stage least square model in the previous section. Also, the quantiles with the highest and lowest coefficients also slightly changed. The results for the public sector were still the same. Returns to schooling in the public sector were the highest at the lowest point of the wage distribution, and then became lower as the point of wage distribution gets lower. However, the lowest marginal return to schooling was not at the 90^{th} quantile but the 75^{th} quantile. Also, the marginal return to schooling at the 10^{th} quantile was 5.5% point higher than that from the ordinary quantile regression, which was 21.3%. The marginal return to schooling at the 75^{th} conditional quantile was 16.2%. Return to schooling in the formal private sector was the highest at the 50^{th} quantile and the lowest at the 10^{th} quantile. The marginal return to schooling in the formal private sector increased as the quantile became larger and then decreased after the 50^{th} quantile. The marginal return to schooling in the informal private sector was the lowest at the 90^{th} conditional quantile and the highest at the 25^{th} quantile. The marginal returns to schooling in the informal private sector were 14.1% at the 25^{th} quantile and 10.2% at the 90^{th} quantile. The

marginal return to schooling became lower as the quantile increased between the 25th and 90th quantiles. The difference between the highest and lowest marginal returns to schooling across quantiles in three sectors was similar (around 1% point) in this model. The differences in the public, formal private, and informal private sectors were 5.1% point, 4.1% point, and 3.9% point, respectively.

Next, a quantile regression with instrumental variable and selection correction was conducted to control for the endogeneity of both education and inter-sectoral allocation. Estimation results are presented in columns (7), (8), and (9) and the third figure in Figure 8.1. In this model, the Dubin and McFadden variant 1 selection correction term, suggested by Bourguignon et al. (2007), was used. In this model, the marginal return to schooling became higher than the two other models in general, as seen in the two-stage least square with selection correction model in the previous section. Also, the quantiles with the highest and lowest values were also slightly different from the two other models. Return to schooling in the public sector was the largest at the lowest point of the wage distribution (10th quantile here), then decreased as the point of wage distribution became lower. Thus, the lowest marginal return to schooling was at the 90th quantile. In addition, the marginal return to schooling at the 10th quantile was 22.7% while that at the 90th quartile was 17.0%. Return to schooling in the formal private sector was the highest at the 10th quantile and the lowest at the 90th quantile. The marginal return to schooling in the formal private sector decreased as the quantile became larger. The marginal return to schooling in the informal private sector was the lowest at the 90th conditional quantile (13.2%) and the highest at the 90th quantile (21.4%). The marginal returns to schooling in the informal private sector across quantiles were much larger than those of the other two models. The difference between the highest and lowest marginal returns to schooling across quantiles in the

three sectors was higher than in the other two models as well, especially for the formal and informal private sectors. The differences between quantiles with the highest and lowest values were 5.8 percentage points in the public sector, 7.6 percentage points in the formal private sector, and 8.2 percentage points in the informal private sector.

Looking at the results of the instrumental variable quantile regression with selection correction, the patterns in the returns to schooling in the three sectors were quite different and did not overlap at most of the quantiles. They were especially different between formal private and informal private sectors and public and informal private sectors. Marginal returns to schooling were the highest at the 10^{th} quantile, which was the lowest quantile, and the lowest at the 90^{th} quantile, the highest quantile, across sectors.

Marginal returns to schooling at different quantiles by sector and gender are shown in Table 8.33 and Figure 8.2. The distributions of coefficients at different quantiles for female and male workers were quite different. The coefficient of years of schooling at different quantiles for the public sector continued to have a decreasing trend for males, but the highest value at the 10^{th} quantile was much higher than that for the total sample. The value at the 10^{th} quantile for the public sector was 33.4% for males and 22.7% for the total sample. Also, the distribution of coefficients for females in the public sector tended to increase—the opposite for the total sample and males. The distribution of the coefficients of years of schooling for males in the formal private sector was similar to that for total, but with a steeper curve. In contrast, the distribution of the coefficients for females in the formal private sector was very different: the curve is concave and the middle parts are lower than the 10^{th} and 90^{th} quantiles. The coefficient of years of schooling at different quantiles for the informal private sector had a slightly increasing trend for males—the opposite for the total sample and for females. Also the conditional distributions in

three sectors tended to overlap at lower quantiles for females but not for males. The return to schooling in the formal private sector was the highest at most quantiles for male. Return to schooling in the public sector was the highest at most quantiles for female.

Tables 8.34-8.37 and Figures 8.7-8.10 display the marginal return to schooling at different quantiles by sector and gender for each population group, estimated by the instrumental variable quantile regression with DMF variant 1 selection correction. We observed the difference in trends by gender and race in the instrumental variable quantile regression. Again, the estimations in the informal sector for Whites and Asians are not available. Estimation by race shows that the formal private sectors had the greatest return to schooling at any quantiles except for Black female and Asians but the distributions of return to schooling in the public and informal private sectors were similar and/or sometimes overlapped for Black total, Black males, Colored males. Black females and Colored females had different trends: the public sector had the highest return to schooling at most quantiles for Blacks, and the public sector has the lowest return to schooling at most quantiles for Coloreds.

8.7.2. Results of the estimation of piecewise spline function

Table 8.38 shows the estimated marginal return to schooling, using the piecewise linear spline function. This model controls only for sample selection into sectors (by DMF variant 1 selection correction), not endogeneity of education. This is because the model has several variables for education and it is difficult to control for the endogeneity of different levels of education. In general, convex marginal returns to schoolings were observed, suggesting that schooling has more impact on earnings as the level of education increases. This is especially true for the formal private and informal private sectors.

Columns (1), (2), and (3) show the marginal return to schooling by level of education in

the public, formal private, and informal private sectors. The marginal return to schooling was the highest for upper secondary education, followed by that for higher education, in the public sector. The marginal return to schooling was the lowest for lower secondary education in the public sector. The marginal returns to upper secondary and higher education were much higher than the marginal returns to primary and lower secondary education. The marginal returns to primary, lower secondary, upper secondary, and higher education in the public sector were 3.8%, 2.0%, 18.2%, and 14.1%, respectively. The marginal return to schooling was the highest for higher education in the formal and informal private sectors, followed by that for upper secondary education. The marginal returns to schooling were the lowest for primary education, followed by that for lower secondary education, in the formal and informal private sectors. In both the formal and informal private sectors, the marginal returns to higher education were much higher than those for other levels of education. The marginal returns to schooling in the formal private sector were 5.5% for primary, 7.1% for lower secondary, 10.4% for upper secondary, and 35.2% for higher education. The marginal returns to primary, lower secondary, upper secondary, and higher education in the informal private sector were 2.8%, 7.9%, 9.2%, and 40.8%, respectively. It is interesting to see that the marginal returns to higher education were the highest in the informal sector.

The marginal returns to different levels of education had similar characteristics for both male and female, with some exceptions. For example, the marginal return to primary education in the public sector was lower than that to lower secondary for males. The marginal return to upper secondary education in the public sector was lower than that to higher education for males. Also, the marginal return to lower secondary education in the informal sector was higher than that to upper secondary education, although it was opposite for the total sample. The marginal

returns to lower secondary and higher education in the public sector were higher for males than for females, while the marginal returns to primary and upper secondary education in the public sector were higher for females than for males. The marginal returns to different level of education in the formal private sector were higher for females than for males, except for primary education. In contrast, the marginal returns to primary and upper secondary education in the informal private sector were higher for males, while the marginal returns to lower secondary and higher education were higher for females.

Table 8.30 show the estimated marginal returns to schooling using the piecewise linear spline function by sector and gender for different population groups. The results for Blacks are similar to the sample of all races. For Coloreds, the marginal return to higher education is the highest in the public sector for Colored total, and Colored females, while the marginal return to upper secondary is the highest in the public sector for Colored males. Also the marginal return to primary schooling is the highest in the informal private sector for Colored total and Colored female while the marginal return to upper secondary is the highest for the Colored male. The trends for Whites and Asians are very different. The marginal return to schooling is the highest for higher education in the public sector for total and female Whites, while it is the highest for lower secondary for male Whites[4]. The marginal return to schooling is the highest for lower secondary education in the formal private sector for total and male Whites; it is the highest for higher education for female Whites. The most of coefficients are not statistically significant for Asians but the marginal return to higher education is the highest in the formal private sector for total Asians, while the marginal return to primary education is the highest in the formal private

[4] This is among the statistically significant coefficients.

sector for female Asians.

8.8. Interpretation and Policy Recommendations

The analyses above show the substantial differential in returns to schooling between the formal and informal private sectors.

Three concrete possible reasons can explain the difference in return to schooling between the formal and the informal private sectors (these reasons are also partly discussed in Nakamuro & Yamasaki, 2012). First, the quality of education that the workers in the informal private sector received could be lower than that which the workers in the formal private sector attained. The second possible reason is that the schooling is being rewarded differently in each sector, perhaps because education is not being fully considered in the informal sector (due to, for example, discrimination) or because education has a lower productivity in the informal sector. If valid, this reason would prove the dual labor market theory. Third, return to schooling in the informal sector is lower because workers do not have to pay tax.

The first reason is related to the education policy in South Africa. In South Africa, access to and quality of education are different by race.[5] Blacks and Coloreds tend to be excluded from the formal private sector and are much more likely to work in the informal private sector. The researcher first thought the difference in returns to schooling between the formal and informal private sectors could be due to the difference in the quality of education between races. However, considering there was still a return to schooling differential among workers of a single race across sectors, this could not fully explain the difference in return to schooling. Another possible reason is that one population group is also grouped into elites and non-elites. The two groups go to

[5]As mentioned in Chapter 5, many schools with students from a single race remained even after the abolishment of the apartheid.

different types of schools with different quality education. One possible policy implication is that the government should provide training and non-formal schooling for workers in the informal sector to complement the low quality of education they may have attained so that they will be become more productive. However, it is also important to investigate further the quality of education received by informal workers, specifically their exact reading and numeracy ability, problem-solving skills, and professional knowledge and skills.

The second possible reason for the differences in return to schooling between the formal and informal private sectors concerns the labor market policy in South Africa. Trade unions are powerful and assertive in the formal private sector in South Africa. Many point out that the formal private sector labor market is distorted because of strong trade unions. The wage is set higher than actual productivity (e.g., Kingdon & Knight, 2007). Therefore, returns to schooling could be high without reflecting productivity. However, it is also possible that the informal private sector, which is thought to be an imperfect labor market, does not properly appreciate schooling or does not efficiently utilize schooling for productivity purposes. Insofar as moving from the informal to the formal sector may allow informal sectgor firms to upgrade their technology and to have access to public infracstructure and to public programs supporting private sector development, efforts to provide incentives to informal sector firms to become formal may pay off in the form of greater productivity, for both the firms themselves as well as the government. Moreover, the government should remove excessively complicated regulations on the labor market and business that lead some firms to move to the informal sector. Policymakers should consider drastically reforming labor market policies and regulations in South Africa.

Third, it is possible that the return to schooling is lower because workers do not pay

taxes. With this assumption, net earnings could be similar between formal and informal sectors. This, however, is not necessarily the case in South Africa, where even formal sector workers tend to avoid income taxes. Although it may apply to the issue of business or sales taxes, it may not be as significant in the case of income taxes. Still, the research in this dissertation sought to ascertain the role of this issue. Although the survey upon which this dissertation is based does not have data on taxes and relies on pre-tax earnings, a preliminary analysis was done, adjusting the wages of the formal sector by the actual tax rates prevailing in South Africa. However, the wage differentials among sectors still were significant and there was no major difference in the results in terms of rates of returns to education.

Lastly, there is the issue of how public sector wages and employment may distort wages and employment in other sectors, including the absorption of highly-educated workers. The issue of the public sector could be analyzed in greater detail in future studies.

8.9. Qualifications and Implications for Future Studies

Conducting this research revealed some concerns about data and methodologies. First, it is possible that workers underreport their earnings and this could be more prevalent in a certain sector. Kuepie et al. (2009) suggest that measurement errors might appear in reported earnings, especially for non-salaried workers in the informal sector. Second, there could be measurement errors in reported years of schooling as well. Third, it is possible that some workers with specific characteristics do not report their earnings at all. Fourth, the validity of the instrument, spouse years of schooling is still a contentious question although it has been used by labor economists many times. Future studies should analyze the validity of the instrument as well as looking for better identification strategies.

Although it is out of the scope of this study, the issue of wages and employment in the

public sector could be analyzed in greater detail in future studies. It is also important to investigate further the details of the quality of education for workers in the formal and informal sectors to find out the reason for the difference in return to schooling between the two sectors.

8.10. Summary and Conclusion

This chapter attempted to clarify the impact of education on earnings and the probability of employment for workers in the formal and informal sectors in South Africa, using the Labor Force Survey in 2007. The study controlled for endogeneity of education and sector sample selection bias and considered the heterogeneity in returns to schooling. It conducted careful analyses using different kinds of methodologies such as ordinary least square, two-stage least square, two-step selection model, and two-stages least squares (2SLS) with the two-step selection model. It also used the wage decomposition method to look at the effect of returns to schooling and workers' average education endowment on the wage difference between different sectors using 2SLS. Also, the quantile regression and piecewise linear spline function were used to explore the heterogeneity in returns to schooling. In the quantile regression analyses, the instrumental variable quantile regression with selection correction was also used.

After controlling for the endogeneity of schooling and sector and labor participation sample selection bias by the 2SLS with Durbin and Mcfadden selection correction, the marginal return to schooling in the informal sector is 16.8% for total, 14.3% for males and 18.7% for females, while that in the public sector is 18.1% for total, 14.5% for males and 27.1% for females and that in the formal private sector is 25.1% for total, 24.2% for males, and 25.35% for females. The significant difference in return to schooling between the formal and informal private sectors is observed across race and gender except for Whites and Asians.

The finding is further corroborated in the decomposition of wage gaps and quantile

regressions which also corrected for the endogeneity of schooling and sample selection bias as well as 2SLS/selection correction models with additional dummies and with different definition of informal sector.

The study confirmed the dual labor market theory: the formal and informal labor markets appear to be segmented. The possible policy implications drawn from the analysis are 1) improving the quality of schools from which informal sector workers graduate, 2) providing training and non-formal schooling for workers in the informal sector, 3) removing complicated regulations in the labor market and seeking to provide incentives for informal sector firms to become formal.

Lastly, future studies should consider dealing with measurement errors in reported earnings and reported schooling. Also the issue of the return to schooling in the public sector and the quality of education of workers in different sectors should be examined in detail.

Table 8.1 *OLS Estimation of the Effect of Years of Schooling on Earnings by Sector, Gender, and Race, All Groups*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.137***	0.096***	0.044***	0.114***	0.095***	0.053***	0.160***	0.100***	0.039***
	(0.005)	(0.003)	(0.004)	(0.007)	(0.003)	(0.008)	(0.008)	(0.005)	(0.005)
N	2728	9548	3054	1320	6107	1170	1408	3441	1884
White	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
years of schooling	0.185***	0.164***	0.094	0.184***	0.156***	0.055	0.179***	0.177***	0.193
	(0.023)	(0.018)	(0.116)	(0.034)	(0.023)	(0.185)	(0.028)	(0.029)	(0.018)
N	175	571	17	86	309	9	89	262	8
Black	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
years of schooling	0.136***	0.088***	0.042***	0.111***	0.087***	0.053***	0.158***	0.092***	0.035***
	(0.006)	(0.003)	(0.005)	(0.007)	(0.004)	(0.009)	(0.010)	(0.006)	(0.006)
N	2029	6630	2514	949	4377	966	1080	2253	1548
Colored	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)
years of schooling	0.138***	0.107***	0.057***	0.116***	0.106***	0.054**	0.160***	0.109***	0.054***
	(0.012)	(0.005)	(0.011)	(0.018)	(0.007)	(0.018)	(0.016)	(0.009)	(0.014)
N	463	2079	512	243	1246	188	220	833	324
Asian	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(46)	(47)	(48)	(49)	(50)	(51)	(52)	(53)	(54)
years of schooling	0.171***	0.147***	0.048	0.148***	0.165***	0.246	0.278	0.106***	-
	(0.038)	(0.026)	(0.131)	(0.036)	(0.036)	-	(0.139)	(0.030)	-
N	61	268	11	42	175	7	19	93	4

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses. Highlighted estimates are better than 2SLS (See Tables 8.3 and 8.4).

Table 8.2 *OLS Estimation of the Effect of Years of Schooling on Earnings by Sector, Gender and Race, Married People*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.123***	0.102***	0.052***	0.106***	0.100***	0.052***	0.148***	0.112***	0.051***
	(0.007)	(0.004)	(0.007)	(0.008)	(0.004)	(0.011)	(0.012)	(0.007)	(0.008)
N	1583	4716	1326	907	3371	521	676	1345	805
White	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
years of schooling	0.170***	0.168***	-0.071	0.179***	0.166***	-0.20	0.155***	0.170***	0.733
	(0.026)	(0.022)	(0.090)	(0.039)	(0.028)	.	(0.032)	(0.035)	.
N	131	396	11	64	215	6	67	181	5
Black	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
years of schooling	0.119***	0.092***	0.049***	0.100***	0.088***	0.051***	0.144***	0.107***	0.048***
	(0.008)	(0.005)	(0.007)	(0.008)	(0.005)	(0.013)	(0.015)	(0.010)	(0.009)
N	1116	3062	1022	634	2312	408	482	750	614
Colored	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
years of schooling	0.130***	0.118***	0.059***	0.113***	0.120***	0.051*	0.153***	0.112***	0.059***
	(0.015)	(0.007)	(0.014)	(0.019)	(0.009)	(0.025)	(0.023)	(0.012)	(0.017)
N	292	1078	288	178	722	103	114	356	185
Asian	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)
years of schooling	0.171***	0.176***	0.12	0.155***	0.180***	0.12	0.268	0.170***	-
	(0.043)	(0.033)	-	(0.038)	(0.044)	-	(0.160)	(0.034)	-
N	44	180	5	31	122	4	13	58	

Note: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses.

Table 8.3 *Tests for 2SLS by Sector and Gender, for All Races*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(Test of endogeneity)	F=22.59	F=170.01	F=23.12	F=12.09	F=114.75	F=5.06	F=10.74	F=45.86	F=19.54
Wu-Hausman F test	p=0.0000	p=0.0000	p=0.0000	p=0.0053	p=0.0000	p=0.0250	p=0.0011	p=0.0000	p=0.0000
(Test of weak instrument)	F=329.07	F=939.69	F=203.86	F=217.51	F=653.21	F=72.52	F=121.68	F=428.84	F=130.38
Test of excluded instrument	p=0.0000	p=0.0000	p=0.0000	p=0.0000	p=0.0000	p=0.0000	p=0.0000	p=0.0000	p=0.0000
N	1312	3963	1140	792	2805	454	520	1158	686

Notes: The Wu-Hausman F-test is based on the estimation with normal standard errors, not robust standard errors. The F-statistics for the test of excluded instrument is same as the Kleibergen-Paap rk Wald F-statistic used for the weak identification test for the IV-GMM.

Table 8.4 *Tests for 2SLS by Sector and Gender, by Race*

White	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(Test of endogeneity)	F=2.77	F=13.52	-	F=0.03	F=4.66	-	F=5.88	F=11.24	-
Wu-Hausman F test	p=0.0986	p=0.0003	-	p=0.8725	p=0.0320	-	p=0.0185	p=0.0010	-
(Test of weak instrument)	F=66.83	F=106.34	-	F=41.23	F=77.57	-	F=28.48	F=34.81	-
Test of excluded instrument	p=0.0000	p=0.0000	-	p=0.0000	p=0.0000	-	p=0.0000	p=0.0000	-
N	127	383	11	63	212	6	64	171	5
Black	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
(Test of endogeneity)	F=14.58	F=94.92	F=16.03	F=4.73	F=73.86	F=3.10	F=9.33	F=18.57	F=14.22
Wu-Hausman F test	p=0.0001	p=0.0000	p=0.0001	p=0.0302	p=0.0000	p=0.0793	p=0.0024	p=0.0000	p=0.0002
(Test of weak instrument)	F=186.99	F=533.56	F=158.39	F=129.82	F=399.71	F=59.22	F=61.05	F=133.14	F=98.90
Test of excluded instrument	p=0.0000	p=0.0000	p=0.0000	p=0.0000	p=0.0000	p=0.0000	p=0.0000	p=0.0000	p=0.0000
N	865	2377	853	524	1782	345	341	595	508

Notes: The Wu-Hausman F test is based on the estimation with normal standard errors not robust standard errors. The F-statistics for the test of excluded instrument is same as the Kleibergen-Paap rk Wald F statistic used for the weak identification test for the IV-GMM.

Table 8.4 (continued)

Colored	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
(Test of endogeneity)	F=8.28	F=61.39	F=7.96	F=12.05	F=40.75	F=3.02	F=0.0076	F=17.94	F=4.86
Wu-Hausman F test	p=0.0043	p=0.0000	p=0.0052	p=0.0007	p=0.0000	p=0.0858	p=0.9309	p=0.0000	p=0.0289
(Test of weak instrument)	F=98.11	F=229.79	F=41.96	F=54.44	F=137.77	F=10.16	F=55.50	F=91.62	F=33.46
Test of excluded instrument	p=0.0000	p=0.0000	p=0.0000	p=0.0000	p=0.0000	p=0.0020	p=0.0000	p=0.0000	p=0.0000
N	276	1035	272	174	700	100	102	335	172
Asian	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
(Test of endogeneity)	F=0.03	F=4.08	-	F=0.004	F=1.51	-	F=0.04	F=2.63	-
Wu-Hausman F test	p=0.8598	p=0.0449	-	p=0.9530	p=0.2212	-	p=0.8534	p=0.1110	-
(Test of weak instrument)	F=10.97	F=24.51	-	F=7.92	F=10.96	-	F=8.35	F=16.61	-
Test of excluded instrument	p=0.0021	p=0.0000	-	p=0.0094	p=0.0013	-	p=0.0202	p=0.0002	-
N	44	168	4	31	111	3	13	57	-

Notes: The Wu-Hausman F test is based on the estimation with normal standard errors not robust standard errors. The F-statistics for the test of excluded instrument is same as the Kleibergen-Paap rk Wald F statistic used for the weak identification test for the IV-GMM.

Table 8.5 *2SLS Estimation by Sector, Gender, and Race*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.170***	0.195***	0.123***	0.143***	0.190***	0.117***	0.217***	0.197***	0.125***
	(0.014)	(0.009)	(0.017)	(0.016)	(0.010)	(0.028)	(0.023)	(0.016)	(0.021)
N	1312	3963	1140	792	2805	454	520	1158	686
White	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
years of schooling	0.224***	0.260***	-0.118	0.183**	0.238***	-0.199	0.275**	0.299***	0.733
	(0.054)	(0.037)	(0.099)	(0.062)	(0.044)	-	(0.088)	(0.066)	-
N	127	383	11	63	212	6	64	171	5
Black	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
years of schooling	0.172***	0.183***	0.115***	0.129***	0.178***	0.109***	0.241***	0.193***	0.116***
	(0.018)	(0.011)	(0.019)	(0.019)	(0.013)	(0.031)	(0.031)	(0.022)	(0.023)
N	865	2377	853	524	1782	345	341	595	508
Colored	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
years of schooling	0.174***	0.225***	0.158***	0.181***	0.233***	0.185*	0.152***	0.199***	0.144***
	(0.028)	(0.019)	(0.037)	(0.037)	(0.024)	(0.077)	(0.038)	(0.032)	(0.043)
N	276	1035	272	174	700	100	102	335	172
Asian	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)
years of schooling	0.183**	0.280***	-	0.158*	0.271***	-	0.241	0.273***	-
	(0.062)	(0.056)	-	(0.057)	(0.076)	-	(0.178)	(0.078)	-
N	44	168	4	31	111	3	13	57	-

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses. The estimates in the highlighted columns mean not better than the OLS estimates.

Table 8.6 *Marginal Effects of Spouse and Household Sector Dummies in the Choice Equation*

	Public	Formal private	Informal private	Unemployed/ out of labor force
spouse formal private	-0.003	0.004	0.004	-0.005
	(0.002)	(0.010)	(0.003)	(0.009)
spouse informal private	-0.011***	-0.111***	0.010***	0.111***
	(0.002)	(0.015)	(0.003)	(0.015)
household formal private	-0.016***	0.003	-0.030***	0.042***
	(0.001)	(0.008)	(0.003)	(0.007)
household informal private	0.005***	-0.011	0.041***	-0.036***
	(0.001)	(0.007)	(0.004)	(0.006)
Joint significance test		chi2(12) = 1724.05		
		Prob > chi2 = 0.0000		

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses.

Table 8.7 *Coefficients and Joint Significance of Spouse and Household Sector Dummies in the Wage Equation, All Races*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
spouse formal private	0.119	0.003	0.065	0.072	-0.019	0.143	0.149	0.037	0.067
	(0.080)	(0.038)	(0.079)	(0.105)	(0.046)	(0.172)	(0.126)	(0.062)	(0.096)
spouse informal pri.	-0.183**	-0.039	0.024	-0.177*	-0.033	0.043	-0.232*	-0.061	0.048
	(0.060)	(0.036)	(0.050)	(0.074)	(0.040)	(0.079)	(0.099)	(0.079)	(0.065)
household formal pri.	-0.088	0.056	0.049	-0.033	0.076	0.022	-0.134	0.019	0.057
	(0.075)	(0.035)	(0.072)	(0.098)	(0.040)	(0.123)	(0.118)	(0.064)	(0.090)
household informal pri.	-0.024	-0.044*	-0.044	-0.002	-0.065*	-0.112	-0.04	0.021	-0.011
	(0.036)	(0.022)	(0.042)	(0.048)	(0.026)	(0.076)	(0.054)	(0.039)	(0.049)
joint significance test	F=3.64	F= 3.73	F= 1.93	F= 1.86	F= 3.61	F= 0.89	F= 2.20	F=0.87	F=1.58
	p=0.0058	p=0.0049	p=0.1024	p= 0.1159	p= 0.0061	p= 0.4718	p=0.0675	p= 0.4806	p=0.1778
N	1583	4716	1326	907	3371	521	676	1345	805

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses.

Table 8.8 *Coefficients of Spouse and Household Sector Dummies in the Wage Equation, by Race*

White	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
spouse formal private	0.551*	0.339*	0.842	0.356	0.483*	-	1.128**	0.092	-
	(0.233)	(0.168)	(1.024)	(0.260)	(0.224)	-	(0.363)	(0.155)	-
spouse informal pri.	0.109	0.294	1.404	0.072	0.361	-	0.483	0.115	-
	(0.246)	(0.260)	(0.620)	(0.352)	(0.328)	-	(0.278)	(0.265)	-
household formal pri.	-0.446	-0.183	-	-0.221	-0.288	-	-1.103**	0.023	-
	(0.229)	(0.171)	-	(0.265)	(0.222)	-	(0.367)	(0.165)	-
household informal pri.	0.127	0.065	0.861	0.231	0.097	-	-0.049	0.01	-
	(0.114)	(0.089)	(1.146)	(0.165)	(0.133)	-	(0.162)	(0.118)	-
joint significance	F= 1.48	F=1.86	F=1.85	F=0.94	F= 1.70	-	F=2.73	F=0.32	-
	p=0.2139	p=0.1164	p=0.4848	p=0.4460	p=0.1521	-	p=0.0375	p=0.8646	-
N	64	215	11	64	215	6	67	181	5
Black	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
spouse formal private	0.151	-0.047	0.126	0.021	-0.099	0.161	0.23	0.056	0.149
	(0.110)	(0.050)	(0.093)	(0.156)	(0.061)	(0.219)	(0.152)	(0.090)	(0.111)
spouse informal pri.	-0.150*	-0.014	0.045	-0.174*	-0.004	0.063	-0.158	-0.072	0.084
	(0.067)	(0.043)	(0.056)	(0.083)	(0.047)	(0.090)	(0.116)	(0.104)	(0.072)
household formal pri.	-0.113	0.112*	0.003	0.01	0.140**	-0.04	-0.195	0.048	0.018
	(0.101)	(0.044)	(0.084)	(0.139)	(0.050)	(0.145)	(0.145)	(0.091)	(0.103)
household informal pri.	-0.042	-0.076**	-0.069	-0.027	-0.111***	-0.173*	-0.03	0.046	0
	(0.043)	(0.026)	(0.048)	(0.057)	(0.030)	(0.085)	(0.065)	(0.052)	(0.056)
joint significance	F= 2.46	F=4.73	F=1.95	F=1.50	F= 5.67	F=1.19	F=1.17	F=1.42	F=1.98
	p= 0.0438	p=0.0008	p=0.1000	p=0.2000	p=0.0002	p=0.3136	p=0.3247	p= 0.2239	p=0.0958
N	1116	3062	1022	634	2312	408	482	750	614

Table 8.8 (continued)

Colored	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
spouse formal private	-0.065	-0.053	-0.161	0.052	-0.035	-0.039	-0.34	-0.024	-0.16
	(0.148)	(0.066)	(0.147)	(0.204)	(0.082)	(0.304)	(0.209)	(0.107)	(0.180)
spouse informal pri.	-0.394**	-0.167*	-0.083	-0.219	-0.164*	-0.116	-0.477*	-0.111	-0.112
	(0.137)	(0.066)	(0.115)	(0.186)	(0.074)	(0.176)	(0.200)	(0.147)	(0.160)
household formal pri.	0.057	0.046	0.206	-0.035	0.06	0.236	0.308	-0.077	0.163
	(0.138)	(0.064)	(0.137)	(0.195)	(0.075)	(0.224)	(0.171)	(0.114)	(0.178)
household informal pri.	0	0.049	-0.072	0.056	0.086	-0.004	-0.107	-0.038	-0.105
	(0.102)	(0.051)	(0.085)	(0.132)	(0.065)	(0.161)	(0.146)	(0.079)	(0.099)
joint significance	F=2.27	F=1.79	F=0.97	F=0.40	F=1.85	F=0.63	F= 2.79	F=0.57	F= 0.66
	p=0.0623	p=0.1294	p=0.4228	p=0.8058	p=0.1177	p=0.6397	p=0.0301	p= 0.6846	p= 0.6199
N	292	1078	288	178	722	103	114	356	185
Asian	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)
spouse formal private	-0.523	0.071	-	-0.591	0.112	-	-0.264	-0.251	-
	(0.409)	(0.132)	-	(0.355)	(0.151)	-	(0.893)	(0.256)	-
spouse informal pri.		0.115	-		0.003	-		0.331	-
		(0.170)	-		(0.258)	-		(0.181)	-
household formal pri.	0.217	-0.127	-	0.305	-0.181	-	0.264	0.27	-
	(0.287)	(0.136)	-	(0.395)	(0.156)	-	(0.803)	(0.241)	-
household informal pri.	-0.365	-0.12	-	-0.163	-0.172	-	-1.535	-0.012	-
	(0.301)	(0.094)	-	(0.216)	(0.116)	-	(1.404)	(0.191)	-
joint significance	F=0.65	F=0.95	-	F=1.25	F=1.08	-	F= 0.75	F= 1.22	-
	p= 0.5894	p=0.4371	-	p=0.3164	p=0.3677	-	p= 0.5665	p= 0.3149	-
N	44	180	5	31	122	4	13	58	-

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses.

Table 8.9 *Two-Step Selection Correction Models Estimations (Total, Male, and Female)*

Lee	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.059***	0.102***	0.049***	0.054***	0.099***	0.051***	0.067***	0.112***	0.049***
	(0.009)	(0.004)	(0.007)	(0.011)	(0.004)	(0.012)	(0.016)	(0.007)	(0.008)
DMF	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
years of schooling	0.057***	0.102***	0.049***	0.052***	0.099***	0.051***	0.065***	0.112***	0.048***
	(0.009)	(0.004)	(0.007)	(0.011)	(0.004)	(0.012)	(0.017)	(0.007)	(0.008)
DMF variant 1	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
years of schooling	0.051***	0.060***	0.040***	0.048***	0.056***	0.035*	0.057**	0.068***	0.043***
	(0.010)	(0.005)	(0.008)	(0.012)	(0.006)	(0.014)	(0.017)	(0.010)	(0.010)
DMF variant 2	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
years of schooling	0.053***	0.062***	0.040***	0.050***	0.058***	0.035*	0.059***	0.070***	0.043***
	(0.009)	(0.005)	(0.008)	(0.011)	(0.006)	(0.014)	(0.017)	(0.009)	(0.010)
Dahl 1	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)
years of schooling	0.063***	0.102***	0.049***	0.057***	0.099***	0.052***	0.075***	0.111***	0.049***
	(0.009)	(0.004)	(0.007)	(0.011)	(0.004)	(0.012)	(0.016)	(0.008)	(0.008)
Dahl 2	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(46)	(47)	(48)	(49)	(50)	(51)	(52)	(53)	(54)
years of schooling	0.055***	0.067***	0.044***	0.050***	0.065***	0.045**	0.065***	0.067***	0.043***
	(0.009)	(0.005)	(0.008)	(0.011)	(0.006)	(0.014)	(0.017)	(0.010)	(0.010)
N	1583	4716	1326	907	3371	521	676	1345	805

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses. Samples only include married people.

Table 8.10 *Lee Two-Step Selection Correction Estimations*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.059***	0.102***	0.049***	0.054***	0.099***	0.051***	0.067***	0.112***	0.049***
	(0.009)	(0.004)	(0.007)	(0.011)	(0.004)	(0.012)	(0.016)	(0.007)	(0.008)
N	1583	4716	1326	907	3371	521	676	1345	805
White	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
years of schooling	0.151***	0.169***	-0.388	0.155*	0.167***	-0.175	0.134*	0.171***	0.379
	(0.043)	(0.024)	(0.179)	(0.068)	(0.032)	-	(0.062)	(0.037)	-
N	131	396	11	64	215	6	67	181	5
Black	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
years of schooling	0.051***	0.091***	0.047***	0.051***	0.087***	0.051***	0.061**	0.106***	0.046***
	(0.011)	(0.005)	(0.008)	(0.013)	(0.005)	(0.013)	(0.020)	(0.010)	(0.009)
N	1116	3062	1022	634	2312	408	482	750	614
Colored	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
years of schooling	0.055**	0.119***	0.055***	0.054*	0.121***	0.046	0.050	0.112***	0.056**
	(0.020)	(0.007)	(0.014)	(0.025)	(0.009)	(0.027)	(0.032)	(0.012)	(0.017)
N	292	1078	288	178	722	103	114	356	185
Asian	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)
years of schooling	0.067	0.176***	0.121	0.045	0.174***	0.121	0.177	0.162***	-
	(0.047)	(0.033)	-	(0.049)	(0.044)	-	(0.128)	(0.033)	-
N	44	180	5	31	122	4	13	58	-

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses. Samples only include married people.

Table 8.11 *DMF Two-Step Selection Correction Estimations*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.057***	0.102***	0.049***	0.052***	0.099***	0.051***	0.065***	0.112***	0.048***
	(0.009)	(0.004)	(0.007)	(0.011)	(0.004)	(0.012)	(0.017)	(0.007)	(0.008)
N	1583	4716	1326	907	3371	521	676	1345	805
White	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
years of schooling	0.153***	0.170***	-0.414	0.163*	0.168***	-0.174	0.132*	0.171***	0.370
	(0.043)	(0.023)	(0.203)	(0.070)	(0.031)	-	(0.061)	(0.036)	-
N	131	396	11	64	215	6	67	181	5
Black	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
years of schooling	0.048***	0.091***	0.047***	0.049***	0.087***	0.051***	0.058**	0.106***	0.046***
	(0.011)	(0.005)	(0.008)	(0.013)	(0.005)	(0.013)	(0.020)	(0.010)	(0.009)
N	1116	3062	1022	634	2312	408	482	750	614
Colored	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
years of schooling	0.050*	0.119***	0.055***	0.052*	0.121***	0.047	0.043	0.112***	0.055**
	(0.020)	(0.007)	(0.014)	(0.025)	(0.009)	(0.028)	(0.033)	(0.012)	(0.017)
N	292	1078	288	178	722	103	114	356	185
Asian	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)
years of schooling	0.063	0.176***	0.222	0.027	0.176***	0.121	0.187	0.161***	-
	(0.048)	(0.033)	-	(0.051)	(0.044)	-	(0.115)	(0.033)	-
N	44	180	5	31	122	4	13	58	-

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses. Samples only include married people.

Table 8.12 *DMF Variant 1 Two-Step Selection Correction Estimations*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.051***	0.060***	0.040***	0.048***	0.056***	0.035*	0.057**	0.068***	0.043***
	(0.010)	(0.005)	(0.008)	(0.012)	(0.006)	(0.014)	(0.017)	(0.010)	(0.010)
N	1583	4716	1326	907	3371	521	676	1345	805
White	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
years of schooling	0.144**	0.155***	-0.364	0.145	0.158**	-0.169	0.121	0.140*	0.372
	(0.053)	(0.037)	-	(0.085)	(0.048)	-	(0.072)	(0.066)	-
N	131	396	11	64	215	6	67	181	5
Black	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
years of schooling	0.045***	0.053***	0.038***	0.047***	0.049***	0.030*	0.053*	0.062***	0.044***
	(0.011)	(0.006)	(0.009)	(0.014)	(0.007)	(0.015)	(0.021)	(0.013)	(0.011)
N	1116	3062	1022	634	2312	408	482	750	614
Colored	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
years of schooling	0.031	0.068***	0.056**	0.034	0.065***	0.071	0.020	0.069***	0.040
	(0.020)	(0.010)	(0.018)	(0.026)	(0.012)	(0.036)	(0.035)	(0.018)	(0.021)
N	292	1078	288	178	722	103	114	356	185
Asian	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)
years of schooling	0.004	0.070	0.121	-0.057	0.032	0.121	1.439	0.103	-
	(0.056)	(0.044)	-	(0.063)	(0.050)	-	(1.047)	(0.068)	-
N	44	180	5	31	122	4	13	58	-

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses. Samples only include married people.

Table 8.13 *DMF Variant 2 Two-Step Selection Correction Estimations*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.053***	0.062***	0.040***	0.050***	0.058***	0.035*	0.059***	0.070***	0.043***
	(0.009)	(0.005)	(0.008)	(0.011)	(0.006)	(0.014)	(0.017)	(0.009)	(0.010)
N	1583	4716	1326	907	3371	521	676	1345	805
White	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
years of schooling	0.145**	0.157***	-0.253	0.149	0.158***	-0.176	0.125	0.144*	0.378
	(0.051)	(0.037)	-	(0.083)	(0.047)	-	(0.071)	(0.067)	-
N	131	396	11	64	215	6	67	181	5
Black	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
years of schooling	0.046***	0.054***	0.038***	0.048***	0.050***	0.029*	0.055**	0.064***	0.044***
	(0.011)	(0.006)	(0.009)	(0.014)	(0.007)	(0.015)	(0.021)	(0.012)	(0.011)
N	1116	3062	1022	634	2312	408	482	750	614
Colored	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
years of schooling	0.033	0.070***	0.055**	0.037	0.067***	0.070	0.021	0.072***	0.039
	(0.020)	(0.010)	(0.018)	(0.026)	(0.012)	(0.035)	(0.034)	(0.017)	(0.021)
N	292	1078	288	178	722	103	114	356	185
Asian	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)
years of schooling	0.003	0.075	0.121	-0.057	0.038	0.121	1.533	0.108	-
	(0.055)	(0.043)	-	(0.061)	(0.050)	-	(0.939)	(0.067)	-
N	44	180	5	31	122	4	13	58	-

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses. Samples only include married people.

Table 8.14 *Dahl 1 Two-Step Selection Correction Estimations*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.063***	0.102***	0.049***	0.057***	0.099***	0.052***	0.075***	0.111***	0.049***
	(0.009)	(0.004)	(0.007)	(0.011)	(0.004)	(0.012)	(0.016)	(0.008)	(0.008)
N	1583	4716	1326	907	3371	521	676	1345	805
White	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
years of schooling	0.151***	0.169***	-0.351	0.150*	0.165***	-0.199	0.137*	0.170***	0.733
	(0.043)	(0.024)	(0.236)	(0.068)	(0.032)	-	(0.060)	(0.037)	-
N	131	396	11	64	215	6	67	181	5
Black	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
years of schooling	0.057***	0.091***	0.047***	0.054***	0.087***	0.051***	0.071***	0.104***	0.046***
	(0.011)	(0.005)	(0.008)	(0.013)	(0.005)	(0.013)	(0.019)	(0.010)	(0.009)
N	1116	3062	1022	634	2312	408	482	750	614
Colored	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
years of schooling	0.061**	0.119***	0.062***	0.067**	0.121***	0.052*	0.053	0.111***	0.062***
	(0.019)	(0.007)	(0.014)	(0.024)	(0.009)	(0.024)	(0.031)	(0.012)	(0.016)
N	292	1078	288	178	722	103	114	356	185
Asian	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)
years of schooling	0.076	0.175***	0.121	0.052	0.172***	0.121	0.923	0.160***	-
	(0.046)	(0.033)	-	(0.044)	(0.044)	-	(0.694)	(0.031)	-
N	44	180	5	31	122	4	13	58	-

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses. Samples only include married people.

Table 8.15 *Dahl 2 Two-Step Selection Correction Estimations*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.055***	0.067***	0.044***	0.050***	0.065***	0.045**	0.065***	0.067***	0.043***
	(0.009)	(0.005)	(0.008)	(0.011)	(0.006)	(0.014)	(0.017)	(0.010)	(0.010)
N	1583	4716	1326	907	3371	521	676	1345	805
White	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
years of schooling	0.130*	0.152***	0.795	0.066	0.155**	-0.126	0.137	0.153*	0.475
	(0.052)	(0.036)	-	(0.089)	(0.046)	-	(0.085)	(0.066)	-
N	131	396	11	64	215	6	67	181	5
Black	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
years of schooling	0.052***	0.062***	0.042***	0.048***	0.060***	0.036*	0.064**	0.064***	0.044***
	(0.011)	(0.006)	(0.009)	(0.013)	(0.007)	(0.015)	(0.020)	(0.013)	(0.012)
N	1116	3062	1022	634	2312	408	482	750	614
Colored	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
years of schooling	0.039*	0.072***	0.061**	0.042	0.077***	0.070*	0.026	0.060**	0.050*
	(0.020)	(0.010)	(0.018)	(0.025)	(0.013)	(0.035)	(0.035)	(0.018)	(0.023)
N	292	1078	288	178	722	103	114	356	185
Asian	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)
years of schooling	0.016	0.082	0.121	-0.042	0.040	0.121	0.839	0.124	-
	(0.053)	(0.045)	-	(0.063)	(0.052)	-	(1.237)	(0.079)	-
N	44	180	5	31	122	4	13	58	-

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses. Samples only include married people.

Table 8.16 *2SLS Estimation with Different Selection Corrections by Sector and Gender*

Lee	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.185***	0.196***	0.124***	0.146***	0.192***	0.121***	0.275***	0.198***	0.125***
	(0.037)	(0.009)	(0.018)	(0.038)	(0.011)	(0.030)	(0.081)	(0.016)	-0.022
DMF	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
years of schooling	0.189***	0.196***	0.125***	0.148***	0.192***	0.122***	0.280***	0.198***	0.126***
	(0.039)	(0.009)	(0.018)	(0.040)	(0.011)	(0.031)	(0.085)	(0.016)	(0.022)
DMF variant 1	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
years of schooling	0.181***	0.251***	0.168***	0.145***	0.242***	0.143**	0.271**	0.253***	0.187***
	(0.040)	(0.020)	(0.031)	(0.044)	(0.024)	(0.049)	(0.089)	(0.039)	(0.043)
DMF variant 2	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
years of schooling	0.179***	0.245***	0.163***	0.145***	0.237***	0.141**	0.269**	0.245***	0.180***
	(0.040)	(0.019)	(0.030)	(0.043)	(0.023)	(0.049)	(0.088)	(0.037)	(0.041)
Dahl 1	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)
years of schooling	0.180***	0.197***	0.125***	0.146***	0.193***	0.125***	0.253***	0.197***	0.125***
	(0.034)	(0.009)	(0.018)	(0.037)	(0.011)	(0.032)	(0.067)	(0.016)	(0.022)
Dahl 2	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)
years of schooling	0.173***	0.271***	0.169***	0.150***	0.264***	0.145**	0.252***	0.263***	0.196***
	(0.035)	(0.022)	(0.031)	(0.041)	(0.025)	(0.049)	(0.073)	(0.040)	(0.047)
N	1312	3963	1140	792	2805	454	520	1158	686

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses.

Table 8.17 *2SLS Estimation with Lee Selection Correction by Sector, Gender, and Race*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.185***	0.196***	0.124***	0.146***	0.192***	0.121***	0.275***	0.198***	0.125***
	(0.037)	(0.009)	(0.018)	(0.038)	(0.011)	(0.030)	(0.081)	(0.016)	-0.022
N	1312	3963	1140	792	2805	454	520	1158	686
White	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
years of schooling	0.290**	0.270***	-0.377	0.179	0.248***	-0.175	0.433*	0.307***	0.379
	(0.107)	(0.040)	(0.237)	(0.128)	(0.050)	-	(0.177)	(0.071)	-
N	127	383	11	63	212	6	64	171	5
Black	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
years of schooling	0.205***	0.184***	0.117***	0.131**	0.179***	0.111***	0.400*	0.194***	0.118***
	(0.055)	(0.011)	(0.020)	(0.051)	(0.013)	(0.032)	(0.171)	(0.022)	-0.025
N	865	2377	853	524	1782	345	341	595	508
Colored	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
years of schooling	0.143*	0.227***	0.157***	0.204*	0.237***	0.189*	-0.004	0.200***	0.138**
	(0.061)	(0.019)	(0.039)	(0.083)	(0.024)	(0.085)	(0.102)	(0.032)	-0.043
N	276	1035	272	174	700	100	102	335	172
Asian	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)
years of schooling	0.157	0.280***	-	0.115	0.271***	-	0.205	0.274***	-
	(0.086)	(0.057)	-	(0.098)	(0.074)	-	(0.198)	(0.076)	-
N	44	168	4	31	111	3	13	57	-

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses.

Table 8.18 *2SLS Estimation with Durbin-McFadden Selection Correction by Sector, Gender, and Race*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.189***	0.196***	0.125***	0.148***	0.192***	0.122***	0.280***	0.198***	0.126***
	(0.039)	(0.009)	(0.018)	(0.040)	(0.011)	(0.031)	(0.085)	(0.016)	(0.022)
N	1312	3963	1140	792	2805	454	520	1158	686
White	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
years of schooling	0.296**	0.266***	-0.358	0.185	0.245***	-0.174	0.449*	0.303***	0.370
	(0.111)	(0.039)	(0.225)	(0.132)	(0.048)	-	(0.193)	(0.069)	-
N	127	383	11	63	212	6	64	171	5
Black	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
years of schooling	0.208***	0.183***	0.118***	0.130*	0.178***	0.112***	0.407*	0.194***	0.120***
	(0.059)	(0.011)	(0.020)	(0.055)	(0.013)	(0.033)	(0.177)	(0.022)	(0.025)
N	865	2377	853	524	1782	345	341	595	508
Colored	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
years of schooling	0.148*	0.227***	0.158***	0.218*	0.237***	0.188*	-0.020	0.200***	0.138**
	(0.066)	(0.019)	(0.040)	(0.090)	(0.024)	(0.084)	(0.117)	(0.032)	(0.044)
N	276	1035	272	174	700	100	102	335	172
Asian	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)
years of schooling	0.160	0.280***	-	0.113	0.271***	-	0.216	0.273***	-
	(0.084)	(0.057)	-	(0.098)	(0.073)	-	(0.190)	(0.077)	-
N	44	168	4	31	111	3	13	57	-

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses.

Table 8.19 *2SLS Estimation with Durbin-McFadden Variant 1 Selection Correction by Sector, Gender, and Race*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.181***	0.251***	0.168***	0.145***	0.242***	0.143**	0.271**	0.253***	0.187***
	(0.040)	(0.020)	(0.031)	(0.044)	(0.024)	(0.049)	(0.089)	(0.039)	(0.043)
N	1312	3963	1140	792	2805	454	520	1158	686
White	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
years of schooling	0.313*	0.492***	-0.364	0.135	0.389*	-0.169	0.551	0.785*	0.372
	(0.149)	(0.142)	-	(0.166)	(0.166)	-	(0.300)	(0.340)	-
N	127	383	11	63	212	6	64	171	5
Black	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
years of schooling	0.196***	0.230***	0.157***	0.121*	0.230***	0.119*	0.404*	0.222***	0.188***
	(0.058)	(0.025)	(0.035)	(0.055)	(0.029)	(0.053)	(0.185)	(0.050)	(0.050)
N	865	2377	853	524	1782	345	341	595	508
Colored	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
years of schooling	0.135	0.342***	0.210**	0.227*	0.338***	0.229	-0.047	0.338**	0.177*
	(0.071)	(0.059)	(0.066)	(0.105)	(0.070)	(0.121)	(0.120)	(0.106)	(0.086)
N	276	1035	272	174	700	100	102	335	172
Asian	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)
years of schooling	0.179	0.412	0.000	0.116	0.292	0.000	1.443	0.634	-
	(0.129)	(0.327)	-	(0.148)	(0.306)	-	(1.239)	(0.677)	-
N	44	168	4	31	111	3	13	57	-

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses.

Table 8.20 *2SLS Estimation with Durbin-McFadden Variant 2 Selection Correction by Sector, Gender, and Race*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.179***	0.245***	0.163***	0.145***	0.237***	0.141**	0.269**	0.245***	0.180***
	(0.040)	(0.019)	(0.030)	(0.043)	(0.023)	(0.049)	(0.088)	(0.037)	(0.041)
N	1312	3963	1140	792	2805	454	520	1158	686
White	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
years of schooling	0.316*	0.473***	-0.253	0.153	0.379*	-0.176	0.546	0.750*	0.378
	(0.145)	(0.131)	(.)	(0.161)	(0.156)	(.)	(0.293)	(0.315)	(.)
N	127	383	11	63	212	6	64	171	5
Black	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
years of schooling	0.194***	0.223***	0.153***	0.121*	0.223***	0.117*	0.399*	0.212***	0.182***
	(0.057)	(0.023)	(0.034)	(0.054)	(0.028)	(0.052)	(0.181)	(0.046)	(0.048)
N	865	2377	853	524	1782	345	341	595	508
Colored	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
years of schooling	0.133	0.331***	0.208**	0.225*	0.328***	0.236	-0.047	0.326**	0.172*
	(0.070)	(0.055)	(0.065)	(0.104)	(0.066)	(0.123)	(0.118)	(0.099)	(0.083)
N	276	1035	272	174	700	100	102	335	172
Asian	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)
years of schooling	0.174	0.402	-	0.109	0.294	-	1.760	0.625	-
	(0.120)	(0.311)	-	(0.146)	(0.304)	-	(1.210)	(0.681)	-
N	44	168	4	31	111	3	13	57	-

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses.

Table 8.21 *2SLS Estimation with Dahl 1 Selection Correction by Sector, Gender, and Race*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.180***	0.197***	0.125***	0.146***	0.193***	0.125***	0.253***	0.197***	0.125***
	(0.034)	(0.009)	(0.018)	(0.037)	(0.011)	(0.032)	(0.067)	(0.016)	(0.022)
N	1312	3963	1140	792	2805	454	520	1158	686
White	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
years of schooling	0.294*	0.271***	-1.078	0.169	0.247***	-0.199	0.463*	0.306***	0.733
	(0.114)	(0.040)	(2.329)	(0.125)	(0.050)	-	(0.208)	(0.072)	-
N	127	383	11	63	212	6	64	171	5
Black	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
years of schooling	0.199***	0.184***	0.118***	0.131**	0.179***	0.116**	0.337**	0.191***	0.119***
	(0.048)	(0.012)	(0.020)	(0.049)	(0.013)	(0.035)	(0.113)	(0.023)	(0.025)
N	865	2377	853	524	1782	345	341	595	508
Colored	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
years of schooling	0.143*	0.227***	0.154***	0.194*	0.238***	0.201*	-0.001	0.199***	0.138**
	(0.058)	(0.019)	(0.038)	(0.075)	(0.024)	(0.091)	(0.094)	(0.032)	(0.043)
N	276	1035	272	174	700	100	102	335	172
Asian	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)
years of schooling	0.175	0.280***	0.000	0.125	0.271***	0.000	1.946	0.269***	-
	(0.102)	(0.057)	-	(0.092)	(0.075)	-	(2.715)	(0.074)	-
N	44	168	4	31	111	3	13	57	-

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses.

Table 8.22 *2SLS Estimation with Dahl 2 Selection Correction by Sector, Gender, and Race*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.173***	0.271***	0.169***	0.150***	0.264***	0.145**	0.252***	0.263***	0.196***
	(0.035)	(0.022)	(0.031)	(0.041)	(0.025)	(0.049)	(0.073)	(0.040)	(0.047)
N	1312	3963	1140	792	2805	454	520	1158	686
White	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
years of schooling	0.330	0.486***	0.795	0.055	0.347*	-0.126	0.579	0.777*	0.475
	(0.173)	(0.143)	-	(0.224)	(0.144)	-	(0.307)	(0.350)	-
N	127	383	11	63	212	6	64	171	5
Black	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
years of schooling	0.184***	0.255***	0.160***	0.130*	0.258***	0.125*	0.378*	0.232***	0.207***
	(0.048)	(0.026)	(0.035)	(0.053)	(0.031)	(0.052)	(0.147)	(0.051)	(0.060)
N	865	2377	853	524	1782	345	341	595	508
Colored	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
years of schooling	0.135*	0.375***	0.206**	0.216*	0.371***	0.250	-0.030	0.343**	0.188*
	(0.065)	(0.066)	(0.065)	(0.094)	(0.076)	(0.139)	(0.087)	(0.116)	(0.087)
N	276	1035	272	174	700	100	102	335	172
Asian	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)
years of schooling	0.198	0.408	-	0.085	0.306	-	-0.129	0.630	-
	(0.179)	(0.306)	-	(0.159)	(0.326)	-	(1.868)	(0.648)	-
N	44	168	4	31	111	3	13	57	-

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses.

Table 8.23 *2SLS Estimation with DMF Variant 1 Selection Correction by Sector, Gender, Using the Workers in the Informal Sector and/or Workers without Written Contract as the Definition of the Informal Sector*

Total	Total				Male				Female			
	Public	Public, informal	Private, formal	Private, informal	Public	Public, informal	Private, formal	Private, informal	Public	Public, informal	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
years of schooling	0.144***	3.234	0.256***	0.216***	0.128**	0.217	0.247***	0.222***	0.186*	-0.816	0.248***	0.220***
	(0.039)	(13.250)	(0.024)	(0.032)	(0.044)	(0.300)	(0.029)	(0.047)	(0.074)	(1.644)	(0.041)	(0.046)
N	1189	123	3073	2030	731	61	2161	1098	458	62	912	932
White	Total				Male				Female			
	Public	Public, informal	Private, formal	Private, informal	Public	Public, informal	Private, formal	Private, informal	Public	Public, informal	Private, formal	Private, informal
	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)
years of schooling	0.298*	-0.119	0.421**	0.671	-	-	0.347*	0.064	0.788	0.000	0.786	0.520
	(0.146)	(.)	(0.140)	(1.159)	-	-	(0.156)	(0.915)	(0.421)	-	(0.488)	(0.386)
N	124	3	341	53	-	-	187	31	62	2	154	22
Black	Total				Male				Female			
	Public	Public, informal	Private, formal	Private, informal	Public	Public, informal	Private, formal	Private, informal	Public	Public, informal	Private, formal	Private, informal
	(25)	(26)	(27)	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
years of schooling	0.158**	-0.564	0.232***	0.206***	0.134*	-0.147	0.229***	0.200***	0.205*	-0.248	0.224***	0.216***
	(0.051)	(0.729)	(0.028)	(0.037)	(0.058)	(0.440)	(0.035)	(0.050)	(0.098)	(0.216)	(0.053)	(0.057)
N	772	93	1800	1430	480	44	1346	781	292	49	454	649
Colored	Total				Male				Female			
	Public	Public, informal	Private, formal	Private, informal	Public	Public, informal	Private, formal	Private, informal	Public	Public, informal	Private, formal	Private, informal
	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)	(46)	(47)	(48)
years of schooling	0.072	0.580	0.361***	0.250***	0.118	2.455	0.347***	0.320*	-0.054	1.305	0.348**	0.215*
	(0.069)	(0.403)	(0.072)	(0.065)	(0.089)	(1.084)	(0.082)	(0.144)	(0.174)	-	(0.122)	(0.086)
N	254	22	786	521	161	13	530	270	93	9	256	251
Asian	Total				Male				Female			
	Public	Public, informal	Private, formal	Private, informal	Public	Public, informal	Private, formal	Private, informal	Public	Public, informal	Private, formal	Private, informal
	(49)	(50\)	(51)	(52)	(53)	(54)	(55)	(56)	(57)	(58)	(59)	(60)
years of schooling	0.730	0.072	0.382	0.109	-0.748	0.077	0.397	-0.234	4.445**	-	0.396	-1.045
	(1.290)	-	(0.298)	(0.197)	(1.546)	-	(0.338)	(0.216)	(0.059)	-	(0.648)	-
N	39	5	146	26	28	3	98	16	11	2	48	10

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses.

Table 8.24 *Estimation with DMF Variant 1 Selection Correction by Sector, Gender and Race, Using the Workers in the Informal Sector and/or without Any of the Written Contract, Medical Aid/Health Insurance, and Pension/Retirement Fund as the Definition of the Informal Sector*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.184***	0.266***	0.199***	0.152**	0.251***	0.208***	0.271**	0.272***	0.195***
	(0.043)	(0.025)	(0.032)	(0.047)	(0.030)	(0.049)	(0.091)	(0.044)	(0.045)
N	1312	3220	1883	792	2276	983	520	944	900
White	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
years of schooling	0.304*	0.503***	0.290	0.131	0.395*	1.829	0.508	0.832*	0.257
	(0.149)	(0.149)	(0.218)	(0.175)	(0.160)	(2.623)	(0.268)	(0.346)	(0.298)
N	127	364	30	63	201	17	64	163	13
Black	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
years of schooling	0.201**	0.237***	0.196***	0.124*	0.232***	0.189***	0.419*	0.240***	0.201***
	(0.061)	(0.029)	(0.038)	(0.058)	(0.035)	(0.053)	(0.202)	(0.054)	(0.058)
N	865	1882	1348	524	1418	709	341	464	639
Colored	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
years of schooling	0.139	0.373***	0.231***	0.249*	0.357***	0.295*	-0.058	0.377**	0.192*
	(0.077)	(0.075)	(0.064)	(0.117)	(0.086)	(0.137)	(0.122)	(0.138)	(0.084)
N	276	819	488	174	553	247	102	266	241
Asian	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)
years of schooling	0.169	0.582	0.122	0.135	0.429	0.025*	1.395	0.611	0.111
	(0.117)	(0.606)	(0.139)	(0.155)	(0.434)	(0.001)	(1.064)	(1.016)	-
N	44	155	17	31	104	10	13	51	7

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses.

Table 8.25 *2SLS Estimation with DMF Variant 1 Selection Correction by Sector, Gender, Using "Without Any of the Written Contract, Medical Aid/Health Insurance, and Pension/Retirement Fund" as the Definition of the Informal Employment*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.186***	0.267***	0.198***	0.153***	0.252***	0.196***	0.271**	0.276***	0.202***
	(0.041)	(0.024)	(0.037)	(0.045)	(0.029)	(0.052)	(0.084)	(0.044)	(0.060)
N	1312	3220	1563	792	2276	886	520	944	677
White	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
years of schooling	0.291*	0.491***	0.135	0.131	0.383**	-0.326	0.469	0.827*	-1.091
	(0.145)	(0.141)	(1.294)	(0.188)	(0.145)	(4.651)	(0.237)	(0.346)	-
N	127	364	21	63	201	13	64	163	8
Black	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
years of schooling	0.204***	0.242***	0.197***	0.130*	0.235***	0.184**	0.413*	0.250***	0.218**
	(0.057)	(0.030)	(0.045)	(0.053)	(0.036)	(0.056)	(0.177)	(0.056)	(0.082)
N	865	1882	1131	524	1418	643	341	464	488
Colored	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
years of schooling	0.143	0.361***	0.212**	0.247*	0.345***	0.267	-0.037	0.373**	0.174*
	(0.075)	(0.067)	(0.069)	(0.115)	(0.074)	(0.144)	(0.109)	(0.131)	(0.082)
N	276	819	397	174	553	222	102	266	175
Asian	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)
years of schooling	0.169	0.601	0.221	0.125	0.468	-0.117	1.465	0.658	0.119
	(0.114)	(0.632)	(0.316)	(0.137)	(0.456)	-	(1.127)	(1.057)	-
N	44	155	14	31	104	8	13	51	6

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses.

Table 8.26 *2SLS Estimation with DMF Variant 1 Selection Correction with District, Industry, and Occupation Dummies by Sector and Gender, All Races*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.181***	0.251***	0.168***	0.145***	0.242***	0.143**	0.271**	0.253***	0.187***
	(0.040)	(0.020)	(0.031)	(0.044)	(0.024)	(0.049)	(0.089)	(0.039)	(0.043)
N	1312	3963	1140	792	2805	454	520	1158	686
Total with district dummies	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
years of schooling	0.182***	0.233***	0.140***	0.137**	0.226***	0.101*	0.245**	0.230***	0.157**
	(0.044)	(0.021)	(0.035)	(0.047)	(0.025)	(0.049)	(0.086)	(0.039)	(0.051)
N	1312	3963	1140	792	2805	454	520	1158	686
Total with district & industry dummies	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
years of schooling	0.185***	0.223***	0.131***	0.137**	0.203***	0.069	0.253*	0.247***	0.154**
	(0.048)	(0.026)	(0.035)	(0.050)	(0.030)	(0.053)	(0.100)	(0.051)	(0.049)
N	1311	3961	1140	791	2804	454	520	1157	686
Total with district, industry & occupation dummies	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
years of schooling	0.166**	0.204***	0.113**	0.117*	0.186***	0.043	0.227*	0.224***	0.147**
	(0.052)	(0.026)	(0.035)	(0.054)	(0.030)	(0.054)	(0.114)	(0.052)	(0.049)
N	1311	3958	1140	791	2802	454	520	1156	686

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses.

Table 8.27 *2SLS Estimation with DMF Variant 1 Selection Correction with District Dummies by Sector, Gender and Race*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.182***	0.233***	0.140***	0.137**	0.226***	0.101*	0.245**	0.230***	0.157**
	(0.044)	(0.021)	(0.035)	(0.047)	(0.025)	(0.049)	(0.086)	(0.039)	(0.051)
N	1312	3963	1140	792	2805	454	520	1158	686
White	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
years of schooling	0.173	0.418**	0.159	0.113	0.320*	0.123	0.021	0.798*	0.372
	(0.151)	(0.143)	(.)	(0.276)	(0.158)	(.)	(0.150)	(0.396)	(.)
N	127	383	11	63	212	6	64	171	5
Black	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
years of schooling	0.199**	0.209***	0.133***	0.109	0.203***	0.101	0.358*	0.204***	0.156**
	(0.065)	(0.026)	(0.040)	(0.062)	(0.031)	(0.055)	(0.168)	(0.053)	(0.059)
N	865	2377	853	524	1782	345	341	595	508
Colored	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
years of schooling	0.146	0.355***	0.182*	0.227	0.372***	0.137	-0.002	0.339**	0.157
	(0.077)	(0.072)	(0.080)	(0.120)	(0.090)	(0.127)	(0.138)	(0.128)	(0.117)
N	276	1035	272	174	700	100	102	335	172
Asian	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)
years of schooling	0.639	0.561	-	0.615	0.261	-	2.346	1.009	-
	(1.031)	(0.553)	-	(1.588)	(0.278)	-	(.)	(1.192)	-
N	44	168	4	31	111	-	13	57	-

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses.

Table 8.28 *2SLS Estimation with DMF Variant 1 Selection Correction with District and Industry Dummies by Sector and Gender*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.185***	0.223***	0.131***	0.137**	0.203***	0.069	0.253*	0.247***	0.154**
	(0.048)	(0.026)	(0.035)	(0.050)	(0.030)	(0.053)	(0.100)	(0.051)	(0.049)
N	1311	3961	1140	791	2804	454	520	1157	686
White	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
years of schooling	0.067	0.460**	-0.145	0.072	0.327*	0.123	0.024	0.793*	0.372
	(0.117)	(0.146)	(.)	(0.202)	(0.146)	(.)	(0.177)	(0.359)	(.)
N	127	381	11	63	211	6	64	170	5
Black	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
years of schooling	0.200**	0.198***	0.125**	0.105	0.175***	0.088	0.352	0.243***	0.142*
	(0.071)	(0.033)	(0.041)	(0.066)	(0.037)	(0.058)	(0.191)	(0.070)	(0.055)
N	864	2377	853	523	1782	345	341	595	508
Colored	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
years of schooling	0.157	0.368***	0.136	0.231	0.362**	-0.098	-0.009	0.352	0.149
	(0.084)	(0.107)	(0.079)	(0.127)	(0.121)	(0.149)	(0.207)	(0.216)	(0.114)
N	276	1035	272	174	700	100	102	335	172
Asian	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)
years of schooling	4.997	0.519	-	1.706	0.303	-	2.346	0.848	-
	(45.016)	(0.464)	-	(3.504)	(0.338)	-	(.)	(0.828)	-
N	44	168	4	31	111	3	13	57	-

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses.

Table 8.29 *2SLS Estimation with DMF Variant 1 Selection Correction with District, Industry, and Occupation Dummies by Sector, Gender and Race*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.166**	0.204***	0.113**	0.117*	0.186***	0.043	0.227*	0.224***	0.147**
	(0.052)	(0.026)	(0.035)	(0.054)	(0.030)	(0.054)	(0.114)	(0.052)	(0.049)
N	1311	3958	1140	791	2802	454	520	1156	686
White	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
years of schooling	0.061	0.412**	-0.093	-0.114	0.314*	0.123	-0.053	0.849*	0.372
	(0.132)	(0.145)	(.)	(0.213)	(0.158)	(.)	(0.214)	(0.405)	(.)
N	127	381	11	63	211	6	64	170	5
Black	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
years of schooling	0.181*	0.179***	0.100*	0.092	0.164***	0.049	0.389	0.217**	0.142*
	(0.076)	(0.032)	(0.039)	(0.071)	(0.036)	(0.055)	(0.254)	(0.071)	(0.058)
N	864	2376	853	523	1781	345	341	595	508
Colored	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
years of schooling	0.126	0.327**	0.128	0.159	0.307**	-0.071	-0.034	0.317	0.148
	(0.095)	(0.105)	(0.079)	(0.160)	(0.113)	(0.168)	(0.207)	(0.207)	(0.119)
N	276	1035	272	174	700	100	102	335	172
Asian	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)
years of schooling	-3.233	0.627	-	-1.756	0.511	-	0.637	0.989	-
	(15.727)	(0.542)	-	(0.674)	(0.525)	-	(.)	(0.905)	-
N	44	166	4	31	110	3	13	56	-

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses.

Table 8.30 *Blinder-Oaxaca Wage Decomposition: Relative Contribution of Years of Schooling to the Earnings Differential in 2SLS with Heckman Selection Correction, All Races*

	All races		
	Total	Male	Female
	(1)	(2)	(3)
Public/formal private			
Due to endowments	0.502	0.416	0.563
	(0.045)	(0.049)	(0.092)
Due to coefficients	-0.416	-0.615	0.172
	(0.274)	(0.273)	(0.680)
Interaction	-0.115	-0.154	0.046
	(0.076)	(0.070)	(0.181)
N	1272/3960	776/2803	496/1157
Public/informal private			
Due to endowments	0.783	0.533	1.113
	(0.155)	(0.195)	(0.269)
Due to coefficients	0.135	0.166	0.651
	(0.345)	(0.419)	(0.725)
Interaction	0.100	0.106	0.595
	(0.256)	(0.267)	(0.662)
N	1272/1177	776/468	496/709
Formal private/informal private			
Due to endowments	0.299	0.217	0.442
	(0.054)	(0.081)	(0.092)
Due to coefficients	0.540	0.625	0.510
	(0.164)	(0.259)	(0.246)
Interaction	0.207	0.207	0.269
	(0.064)	(0.088)	(0.130)
N	3960/1140	2803/454	1157/686

Table 8.31 *Blinder-Oaxaca Wage Decomposition: Relative Contribution of Years of Schooling to the Earnings Differential in 2SLS with Heckman Selection Correction, by Race*

	White			Black		
	Total	Male	Female	Total	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)
Public/formal private						
Due to endowments	0.232	0.106	0.528	0.562	0.452	0.643
	0.096	0.114	0.237	0.063	0.065	0.144
Due to coefficients	-1.026	-1.583	-3.198	-0.278	-0.641	0.822
	1.406	1.797	2.871	0.354	0.351	0.980
Interaction	-0.054	-0.045	-0.235	-0.091	-0.179	0.266
	0.077	0.069	0.223	0.116	0.100	0.317
Total	-0.848	-1.522	-2.904	0.193	-0.368	1.731
N	125/383	62/212	63/171	833/2376	512/1781	321/595
Public/informal private						
Due to endowments	-	-	-	0.742	0.487	1.115
	-	-	-	0.174	0.210	0.320
Due to coefficients	-	-	-	0.312	0.177	1.586
	-	-	-	0.471	0.514	1.358
Interaction	-	-	-	0.224	0.115	1.395
	-	-	-	0.338	0.333	1.196
Total	-	-	-	1.278	0.779	4.096
N	-	-	-	833/882	512/355	321/527
Formal private/informal private						
Due to endowments	-	-	-	0.236	0.186	0.366
	-	-	-	0.053	0.078	0.097
Due to coefficients	-	-	-	0.469	0.567	0.419
	-	-	-	0.195	0.274	0.313
Interaction	-	-	-	0.146	0.173	0.184
	-	-	-	0.062	0.086	0.138
Total	-	-	-	0.851	0.925	0.969
N	-	-	-	2376/853	1781/345	595/508

Table 8.31 (continued)

	Colored			Asian		
	Total	Male	Female	Total	Male	Female
	(7)	(8)	(9)	(10)	(11)	(12)
Public/formal private						
Due to endowments	0.551	0.522	0.520	0.376	0.334	0.523
	0.100	0.116	0.186	0.187	0.218	0.338
Due to coefficients	-0.819	-0.687	-1.361	-2.481	-3.553	-4.606
	0.480	0.502	1.504	1.831	2.021	2.184
Interaction	-0.257	-0.216	-0.406	-0.257	-0.319	-0.625
	0.153	0.161	0.454	0.209	0.239	0.404
Total	-0.525	-0.381	-1.247	-2.363	-3.538	-4.708
N	271/1035	172/700	99/335	43/166	30/110	13/56
Public/informal private						
Due to endowments	0.593	0.274	0.744	-	-	-
	0.258	0.321	0.408	-	-	-
Due to coefficients	-0.088	0.594	-1.407	-	-	-
	0.581	0.807	1.651	-	-	-
Interaction	-0.063	0.326	-1.263	-	-	-
	0.415	0.445	1.485	-	-	-
Total	0.441	1.194	-1.926	-	-	-
N	271/277	172/102	99/175	-	-	-
Formal private/informal private						
Due to endowments	0.227	0.092	0.349	-	-	-
	0.086	0.117	0.131	-	-	-
Due to coefficients	0.725	1.061	0.439	-	-	-
	0.339	0.650	0.426	-	-	-
Interaction	0.224	0.206	0.201	-	-	-
	0.109	0.144	0.196	-	-	-
Total	1.176	1.360	0.988	-	-	-
N	1035/272	700/100	335/172	-	-	-

Table 8.32 *Marginal Return to Schooling at Different Quantiles by Sector*

	qreg			qreg + IV			qreg + IV + selection		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
10%									
yschool	0.158	0.066	0.052	0.213	0.164	0.137	0.227	0.321	0.214
	(0.009)	(0.004)	(0.010)	(0.011)	(0.006)	(0.010)	(0.012)	(0.008)	(0.012)
25%									
yschool	0.139	0.075	0.036	0.177	0.185	0.141	0.172	0.296	0.193
	(0.008)	(0.004)	(0.006)	(0.008)	(0.005)	(0.008)	(0.009)	(0.006)	(0.009)
50%									
yschool	0.138	0.091	0.048	0.167	0.204	0.129	0.189	0.272	0.164
	(0.005)	(0.003)	(0.004)	(0.007)	(0.005)	(0.007)	(0.008)	(0.006)	(0.008)
75%									
yschool	0.138	0.105	0.041	0.162	0.193	0.109	0.173	0.253	0.181
	(0.005)	(0.003)	(0.005)	(0.008)	(0.005)	(0.008)	(0.009)	(0.006)	(0.009)
90%									
yschool	0.128	0.116	0.052	0.168	0.191	0.102	0.170	0.220	0.132
	(0.007)	(0.004)	(0.006)	(0.010)	(0.006)	(0.010)	(0.012)	(0.007)	(0.011)
Difference between									
the highest and the lowest	0.030	0.049	0.015	0.051	0.041	0.039	0.058	0.076	0.082
OLS/IV/IV+selection									
yschool	0.134	0.095	0.044	0.170	0.195	0.123	0.181	0.251	0.168
	(0.005)	(0.003)	(0.004)	(0.014)	(0.009)	(0.017)	(0.037)	(0.019)	(0.031)
selection correction	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1
N	2728	9553	3054	1312	3963	1140	1312	3963	1140

Note: Robust standard errors are in the parentheses.

Table 8.33 *Marginal Return to Schooling at Different Quantiles by Sector and Gender*

	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
10%									
yschool	0.227	0.321	0.214	0.334	0.319	0.126	0.230	0.288	0.192
	(0.012)	(0.008)	(0.012)	(0.019)	(0.010)	(0.018)	(0.019)	(0.014)	(0.014)
25%									
yschool	0.172	0.296	0.193	0.116	0.316	0.128	0.274	0.223	0.235
	(0.009)	(0.006)	(0.009)	(0.012)	(0.008)	(0.014)	(0.016)	(0.010)	(0.012)
50%									
yschool	0.189	0.272	0.164	0.109	0.280	0.173	0.231	0.241	0.182
	(0.008)	(0.006)	(0.008)	(0.010)	(0.007)	(0.014)	(0.013)	(0.010)	(0.010)
75%									
yschool	0.173	0.253	0.181	0.104	0.248	0.140	0.338	0.221	0.185
	(0.009)	(0.006)	(0.009)	(0.011)	(0.007)	(0.015)	(0.016)	(0.010)	(0.011)
90%									
yschool	0.170	0.220	0.132	0.117	0.191	0.166	0.316	0.261	0.134
	(0.012)	(0.007)	(0.011)	(0.015)	(0.008)	(0.019)	(0.020)	(0.014)	(0.013)
Difference between									
the highest and the lowest	0.058	0.101	0.082	0.229	0.128	0.047	0.108	0.067	0.101
2sls+selection									
yschool	0.181	0.251	0.168	0.145	0.242	0.142	0.271	0.253	0.189
	(0.037)	(0.019)	(0.031)	(0.038)	(0.022)	(0.052)	(0.087)	(0.037)	(0.043)
selection correction	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1
N	1312	3963	1140	792	2805	454	520	1158	686

Note: Robust standard errors are in the parentheses.

Table 8.34 *Marginal Return to Schooling at Different Quantiles by Sector and Gender, Whites*

	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
10%									
yschool	0.179	0.602	-	0.099	0.597	-	0.249	1.436	-
	(0.054)	(0.046)	-	(0.104)	(0.068)	-	(0.072)	(0.113)	-
25%									
yschool	0.083	0.598	-	0.005	0.401	-	0.668	1.072	-
	(0.042)	(0.036)	-	(0.077)	(0.046)	-	(0.072)	(0.067)	-
50%									
yschool	0.159	0.546	-	0.146	0.243	-	0.915	1.150	-
	(0.039)	(0.031)	-	(0.067)	(0.039)	-	(0.082)	(0.066)	-
75%									
yschool	0.202	0.551	-	0.190	0.397	-	0.750	1.288	-
	(0.043)	(0.035)	-	(0.075)	(0.046)	-	(0.076)	(0.077)	-
90%									
yschool	0.377	0.486	-	0.193	0.659	-	0.834	0.539	-
	(0.058)	(0.042)	-	(0.096)	(0.071)	-	(0.118)	(0.057)	-
Difference between									
the highest and the lowest	0.294	0.116	-	0.188	0.257	-	0.667	0.750	-
2sls+selection									
yschool	0.313	0.492	-0.685	0.135	0.389	-0.169	0.551	0.785	0.372
	(0.138)	(0.115)	-	(0.164)	(0.139)	-	(0.265)	(0.285)	-
selection correction	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1
N	127	383	11	63	212	6	64	171	5

Note: Robust standard errors are in the parentheses.

Table 8.35 *Marginal Return to Schooling at Different Quantiles by Sector and Gender, Blacks*

	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
10%									
yschool	0.199	0.235	0.158	0.087	0.308	0.166	0.225	0.128	0.250
	(0.014)	(0.009)	(0.013)	(0.017)	(0.012)	(0.022)	(0.022)	(0.017)	(0.018)
25%									
yschool	0.153	0.270	0.173	0.086	0.302	0.096	0.338	0.093	0.245
	(0.011)	(0.008)	(0.010)	(0.014)	(0.009)	(0.016)	(0.019)	(0.013)	(0.014)
50%									
yschool	0.189	0.245	0.159	0.059	0.264	0.140	0.396	0.237	0.208
	(0.010)	(0.007)	(0.009)	(0.012)	(0.008)	(0.015)	(0.018)	(0.013)	(0.012)
75%									
yschool	0.186	0.251	0.144	0.087	0.240	0.136	0.536	0.234	0.178
	(0.011)	(0.007)	(0.010)	(0.014)	(0.008)	(0.016)	(0.023)	(0.014)	(0.013)
90%									
yschool	0.188	0.214	0.140	0.047	0.190	0.095	0.549	0.241	0.117
	(0.014)	(0.009)	(0.012)	(0.017)	(0.010)	(0.021)	(0.032)	(0.018)	(0.015)
Difference between									
the highest and the lowest	0.046	0.056	0.033	0.041	0.118	0.070	0.171	0.148	0.133
2sls+selection									
yschool	0.196	0.230	0.157	0.121	0.230	0.118	0.404	0.222	0.189
	(0.054)	(0.023)	(0.036)	(0.053)	(0.027)	(0.055)	(0.172)	(0.051)	(0.049)
selection correction	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1
N	865	2377	853	524	1782	345	341	595	508

Note: Robust standard errors are in the parentheses.

Table 8.36 . *Marginal Return to Schooling at Different Quantiles by Sector and Gender, Coloreds*

	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
10%									
yschool	0.316	0.572	0.225	0.559	0.468	0.332	-0.071	0.543	0.144
	(0.034)	(0.025)	(0.023)	(0.062)	(0.027)	(0.047)	(0.054)	(0.039)	(0.027)
25%									
yschool	0.158	0.493	0.243	0.376	0.496	0.457	0.162	0.412	0.218
	(0.021)	(0.017)	(0.019)	(0.036)	(0.021)	(0.041)	(0.041)	(0.024)	(0.023)
50%									
yschool	0.010	0.419	0.178	0.269	0.432	0.200	-0.030	0.356	0.145
	(0.018)	(0.014)	(0.016)	(0.027)	(0.017)	(0.026)	(0.034)	(0.020)	(0.019)
75%									
yschool	0.103	0.368	0.189	0.166	0.378	0.309	-0.139	0.286	0.312
	(0.020)	(0.014)	(0.017)	(0.027)	(0.018)	(0.034)	(0.040)	(0.020)	(0.028)
90%									
yschool	0.109	0.250	0.356	0.167	0.243	0.373	-0.446	0.256	0.083
	(0.027)	(0.015)	(0.030)	(0.036)	(0.018)	(0.047)	(0.076)	(0.025)	(0.025)
Difference between the highest and the lowest	0.306	0.322	0.178	0.393	0.253	0.256	0.608	0.287	0.229
2sls+selection									
yschool	0.135	0.342	0.198	0.227	0.338	0.241	-0.047	0.338	0.182
	(0.056)	(0.054)	(0.065)	(0.081)	(0.067)	(0.140)	(0.121)	(0.092)	(0.088)
selection correction	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1
N	276	1035	272	174	700	100	102	335	172

Note: Robust standard errors are in the parentheses.

Table 8.37 *Marginal Return to Schooling at Different Quantiles by Sector and Gender, Asians*

	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
10%									
yschool	-0.019	0.928	-	0.112	-0.111	-	-	0.582	-
	(0.093)	(0.090)	-	(0.070)	(0.060)	-	-	(0.097)	-
25%									
yschool	0.126	0.330	-	0.083	0.023	-	-	0.474	-
	(0.067)	(0.038)	-	(0.050)	(0.042)	-	-	(0.067)	-
50%									
yschool	0.258	0.095	-	-0.024	0.112	-	-	0.519	-
	(0.064)	(0.030)	-	(0.040)	(0.040)	-	-	(0.062)	-
75%									
yschool	0.023	0.294	-	0.049	0.681	-	-	0.524	-
	(0.067)	(0.036)	-	(0.052)	(0.073)	-	-	(0.066)	-
90%									
yschool	0.085	-0.627	-	0.037	-0.326	-	-	0.705	-
	(0.093)	(0.075)	-	(0.087)	(0.070)	-	-	(0.116)	-
Difference between									
the highest and the lowest	0.277	1.555	-	0.136	0.792	-	-	0.231	-
2sls+selection									
yschool	0.179	0.412	-	0.116	0.292	-	1.443	0.634	-
	(0.164)	(0.341)	-	(0.169)	(0.365)	-	(0.909)	(0.544)	-
selection correction	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1	DMF v1
N	44	168	4	31	111	3	13	57	N/A

Note: Robust standard errors are in the parentheses.

Table 8.38 *Estimation of Piecewise Spline Function, All Races*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
0-6 years of schooling (primary)	0.038**	0.055***	0.029***	0.019	0.058***	0.035*	0.057*	0.046***	0.024*
	(0.014)	(0.006)	(0.008)	(0.016)	(0.007)	(0.014)	(0.024)	(0.010)	(0.010)
7-9 years of schooling (lower sec.)	0.020	0.071***	0.077***	0.039	0.063***	0.066*	0.022	0.086***	0.082***
	(0.025)	(0.010)	(0.017)	(0.028)	(0.012)	(0.029)	(0.044)	(0.016)	(0.021)
10-12 years of schooling (upper sec.)	0.182***	0.104***	0.089***	0.165***	0.086***	0.111**	0.228***	0.131***	0.064*
	(0.021)	(0.010)	(0.023)	(0.026)	(0.012)	(0.041)	(0.033)	(0.016)	(0.029)
13+ (higher)	0.141***	0.352***	0.430***	0.206***	0.333***	0.382*	0.107***	0.391***	0.451***
	(0.024)	(0.028)	(0.102)	(0.037)	(0.034)	(0.159)	(0.030)	(0.052)	(0.134)
N	2728	9553	3054	1320	6109	1170	1408	3444	1884

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses.

Table 8.39 *Estimation of Piecewise Spline Function, By Race*

White	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
0-6 years of schooling (primary)	.	-0.136	-1.940	.	-0.214	.	.	-0.266	.
	.	(0.107)	(0.767)	.	(0.162)	.	.	(0.767)	.
7-9 years of schooling (lower sec.)	0.275	0.377*	0.814	0.393*	0.617*	-0.971	-0.485	0.285	0.415
	(0.205)	(0.160)	(0.310)	(0.192)	(0.280)	(.)	(0.354)	(0.224)	(.)
10-12 years of schooling (upper sec.)	0.151*	0.108**	-0.893*	0.166	0.089	-1.191	0.166	0.113	0.014
	(0.069)	(0.041)	(0.186)	(0.091)	(0.057)	(.)	(0.135)	(0.060)	(.)
13+ (higher)	0.217**	0.304***	0.554	0.180	0.257**	0.292	0.182	0.330**	0.000
	(0.078)	(0.067)	(0.532)	(0.115)	(0.080)	(.)	(0.103)	(0.121)	(.)
N	175	571	17	86	309	9	89	262	8
Black	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
0-6 years of schooling (primary)	0.034*	0.057***	0.023**	0.020	0.061***	0.026	0.047	0.046***	0.018
	(0.016)	(0.007)	(0.009)	(0.018)	(0.008)	(0.015)	(0.028)	(0.011)	(0.011)
7-9 years of schooling (lower sec.)	0.021	0.068***	0.085***	0.047	0.058***	0.074*	0.034	0.090***	0.088***
	(0.030)	(0.012)	(0.019)	(0.034)	(0.014)	(0.033)	(0.053)	(0.020)	(0.024)
10-12 years of schooling (upper sec.)	0.171***	0.109***	0.101***	0.143***	0.087***	0.115*	0.235***	0.138***	0.083**
	(0.026)	(0.012)	(0.026)	(0.035)	(0.015)	(0.045)	(0.041)	(0.019)	(0.031)
13+ (higher)	0.134***	0.469***	0.460**	0.228***	0.429***	0.454*	0.091**	0.571***	0.473**
	(0.028)	(0.037)	(0.142)	(0.044)	(0.042)	(0.217)	(0.033)	(0.066)	(0.161)
N	2029	6635	2514	949	4379	966	1080	2256	1548

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses.

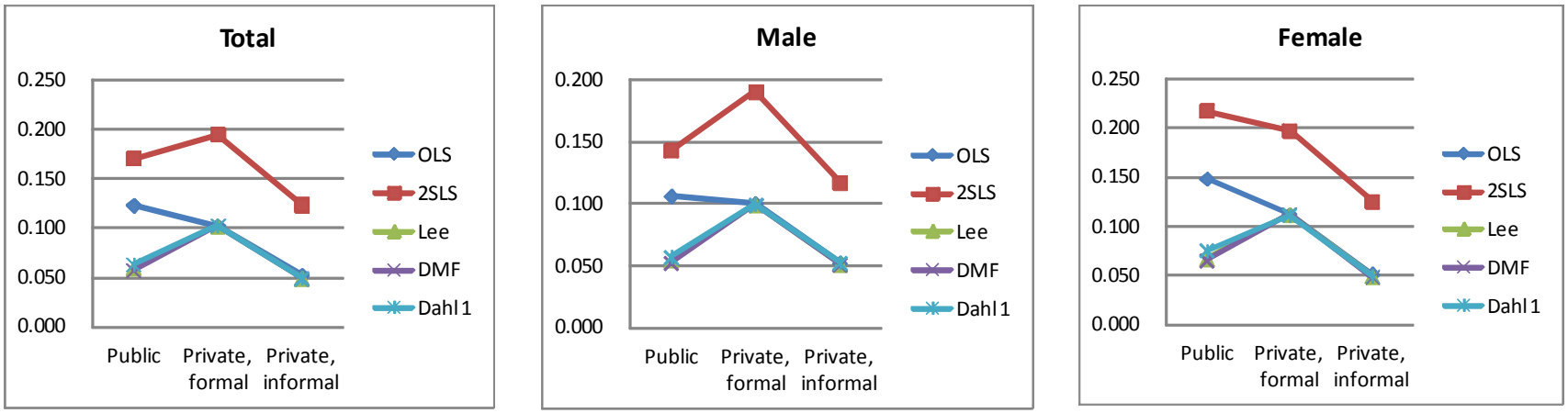
Table 8.39 (continued)

Colored	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
0-6 years of schooling (primary)	0.062*	0.054***	0.061**	0.033	0.054***	0.070	0.094*	0.043	0.054*
	(0.027)	(0.011)	(0.021)	(0.037)	(0.013)	(0.037)	(0.038)	(0.024)	(0.025)
7-9 years of schooling (lower sec.)	-0.014	0.089***	0.043	0.014	0.081***	0.041	-0.002	0.109***	0.030
	(0.046)	(0.018)	(0.041)	(0.055)	(0.024)	(0.069)	(0.080)	(0.026)	(0.053)
10-12 years of schooling (upper sec.)	0.237***	0.113***	0.044	0.275***	0.094**	0.116	0.217***	0.131***	-0.052
	(0.042)	(0.022)	(0.060)	(0.060)	(0.029)	(0.098)	(0.063)	(0.034)	(0.075)
13+ (higher)	0.297***	0.402***	-0.357*	0.270*	0.449**	-0.191	0.259***	0.366**	0.000
	(0.066)	(0.092)	(0.158)	(0.133)	(0.137)	(0.294)	(0.075)	(0.127)	(.)
N	463	2079	512	243	1246	188	220	833	324
Asian	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
0-6 years of schooling (primary)	.	-0.014	.	.	-0.242**	.	.	0.199*	.
	.	(0.074)	.	.	(0.073)	.	.	(0.083)	.
7-9 years of schooling (lower sec.)	0.029	0.038	-0.019	0.246	0.167	-0.328	0.000	-0.268	.
	(0.284)	(0.101)	(.)	(0.145)	(0.123)	(.)	(.)	(0.168)	.
10-12 years of schooling (upper sec.)	0.090	0.067	0.016	-0.010	0.023	0.382	2.600	0.145	.
	(0.126)	(0.043)	(.)	(0.102)	(0.048)	(.)	(1.400)	(0.083)	.
13+ (higher)	-0.020	0.219**	.	-0.037	0.139	.	-0.323	0.175	.
	(0.087)	(0.079)	.	(0.094)	(0.083)	.	(0.377)	(0.159)	.
N	61	268	11	42	175	7	19	93	.

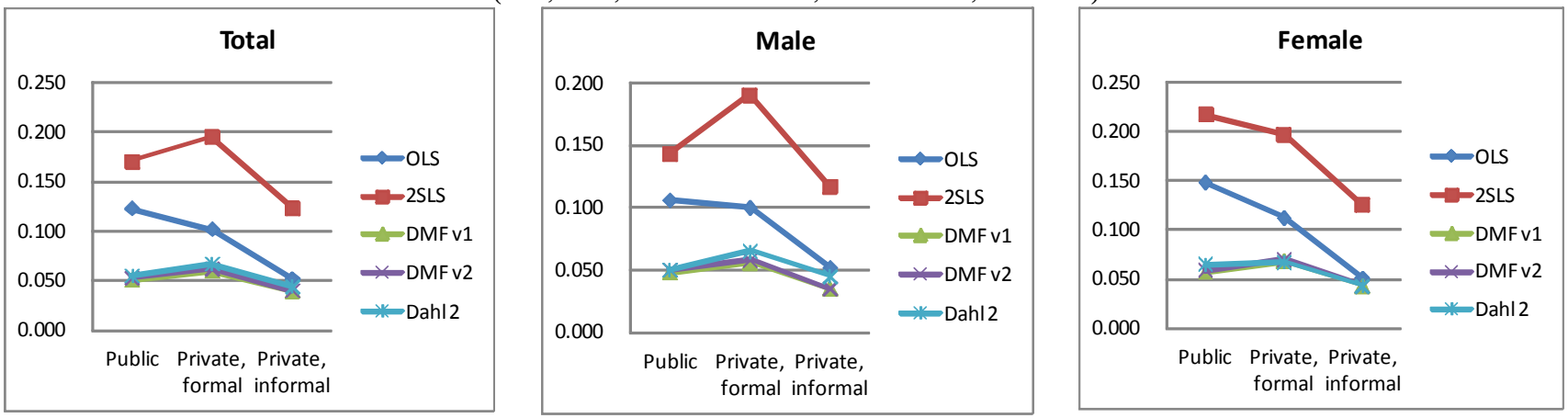
Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in parentheses.

Figure 8.1. The Marginal Return to Schooling in the Public, Formal Private, and Informal Private in Different Models 1, Total

(OLS, 2SLS, and Lee, DMF, and Dahl 1)



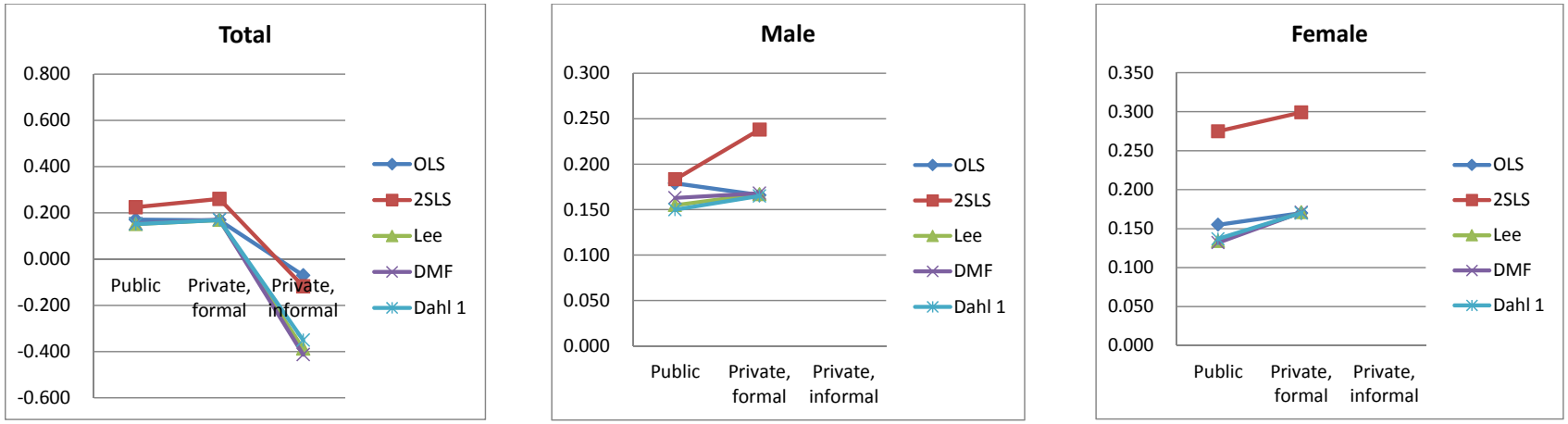
(OLS, 2SLS, and DMF variant 1, DMF variant 2, and Dahl 2)



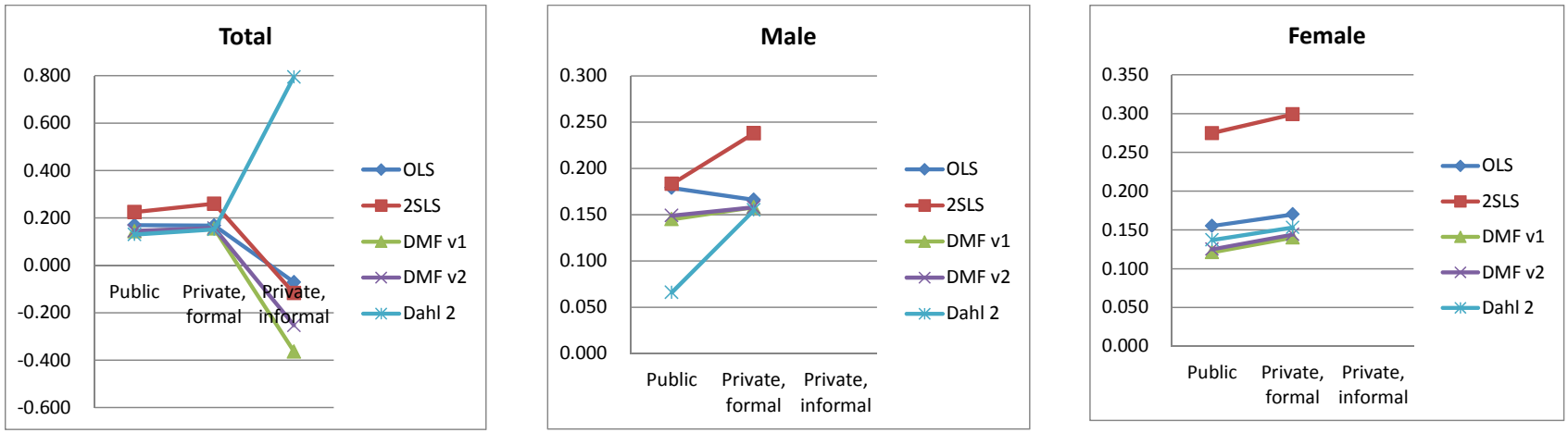
Notes: The sample used the models are only married workers.

Figure 8.2. The Marginal Return to Schooling in the Public, Formal Private, and Informal Private in Different Models 1, Whites

(OLS, 2SLS, and Lee, DMF, and Dahl 1)



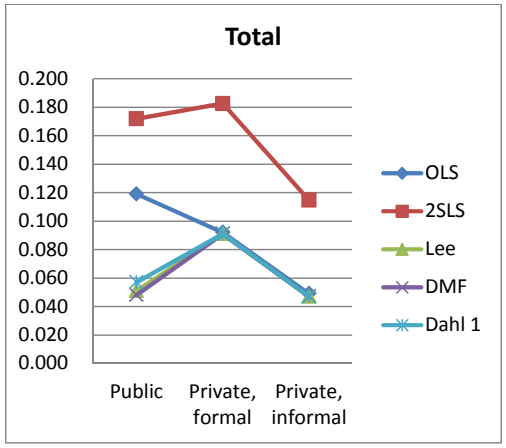
(OLS, 2SLS, and DMF variant 1, DMF variant 2, and Dahl 2)

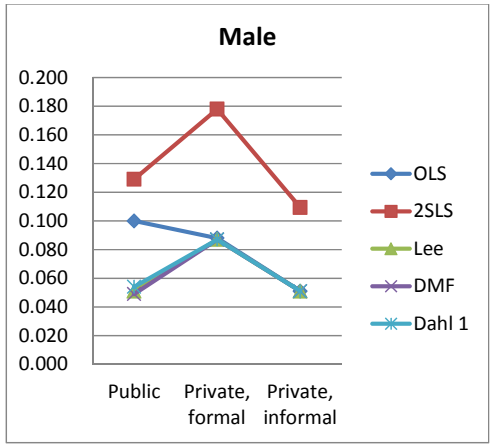


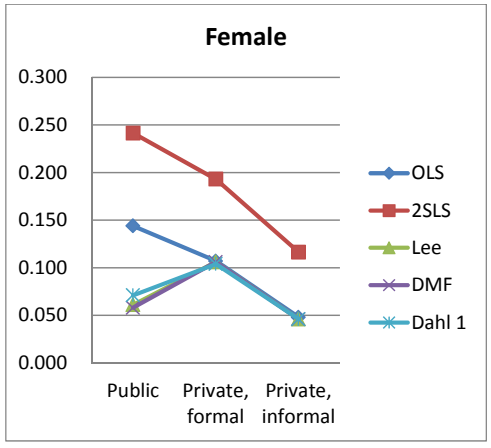
Notes: The sample used the models are only married workers. Not enough sample size is available for the informal sector for Whites.

Figure 8.3. The Marginal Return to Schooling in the Public, Formal Private, and Informal Private in Different Models 1, Blacks

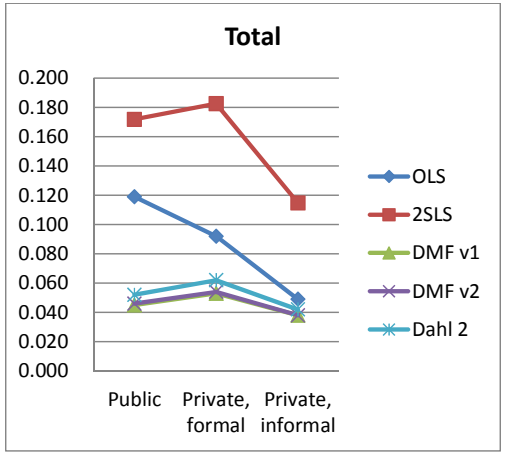
(OLS, 2SLS, and Lee, DMF, and Dahl 1)

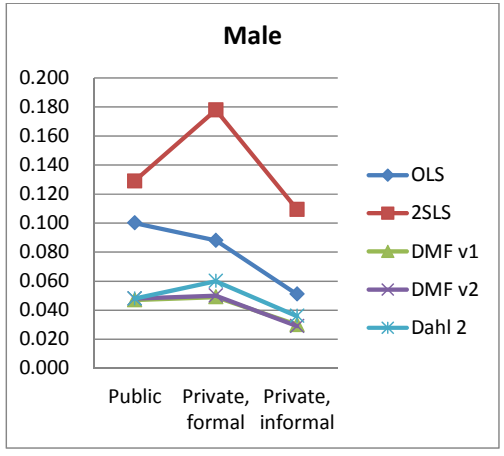


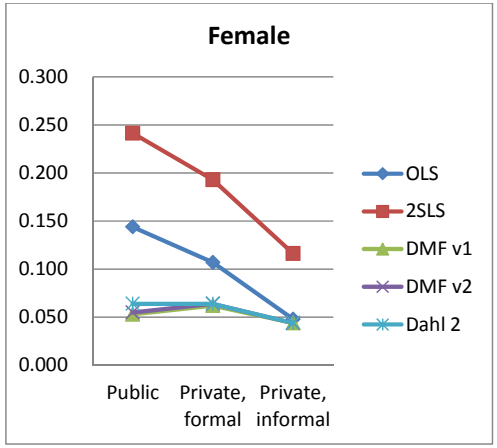




(OLS, 2SLS, and DMF variant 1, DMF variant 2, and Dahl 2)



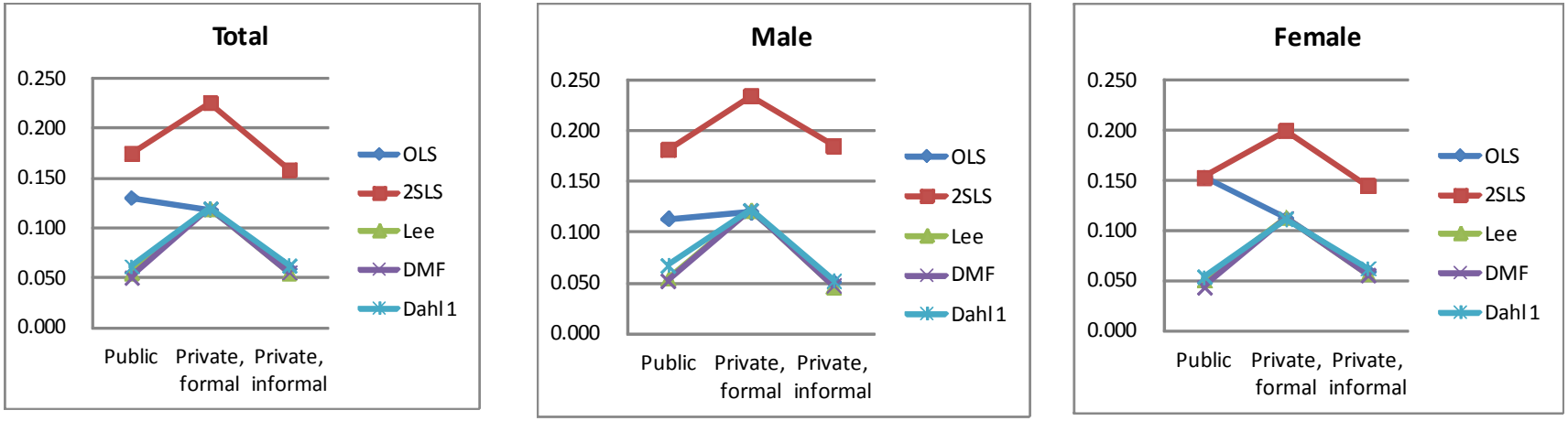




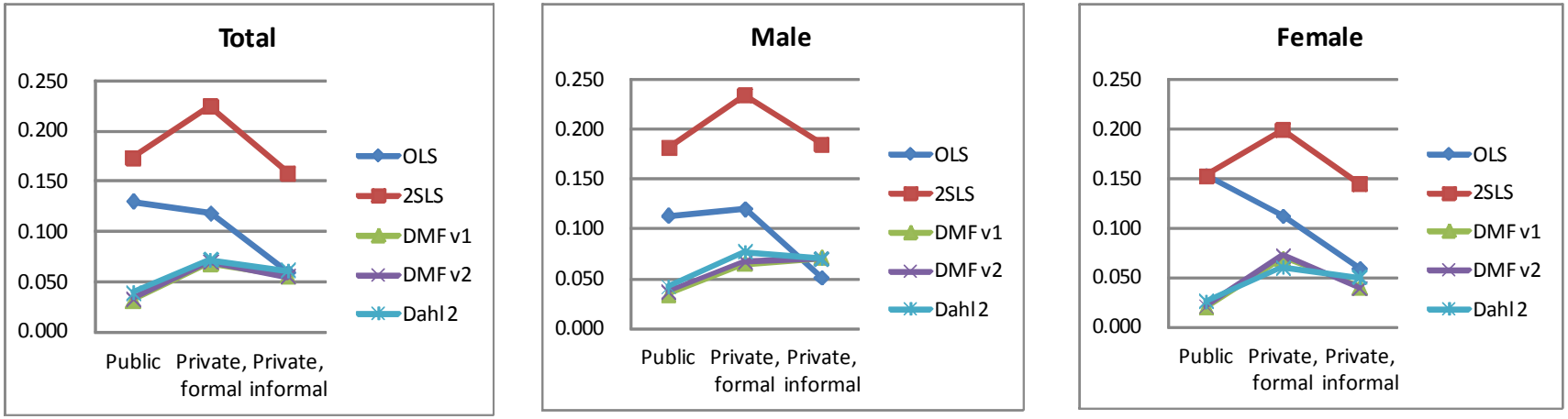
Notes: The sample used the models are only married workers.

Figure 8.4. The Marginal Return to Schooling in the Public, Formal Private, and Informal Private in Different Models 1, Coloreds

(OLS, 2SLS, and Lee, DMF, and Dahl 1)



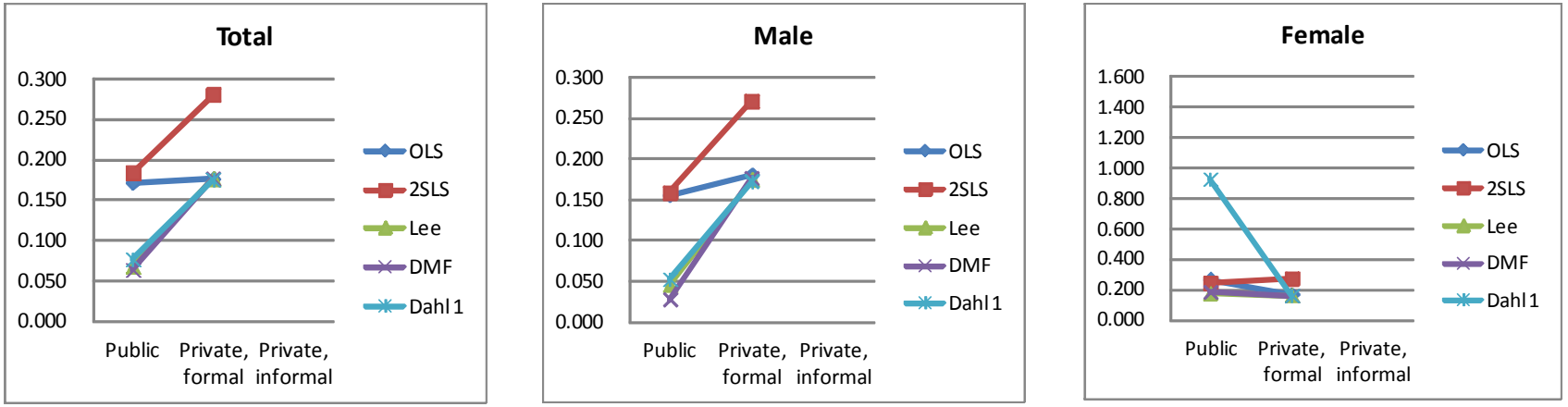
(OLS, 2SLS, and DMF variant 1, DMF variant 2, and Dahl 2)



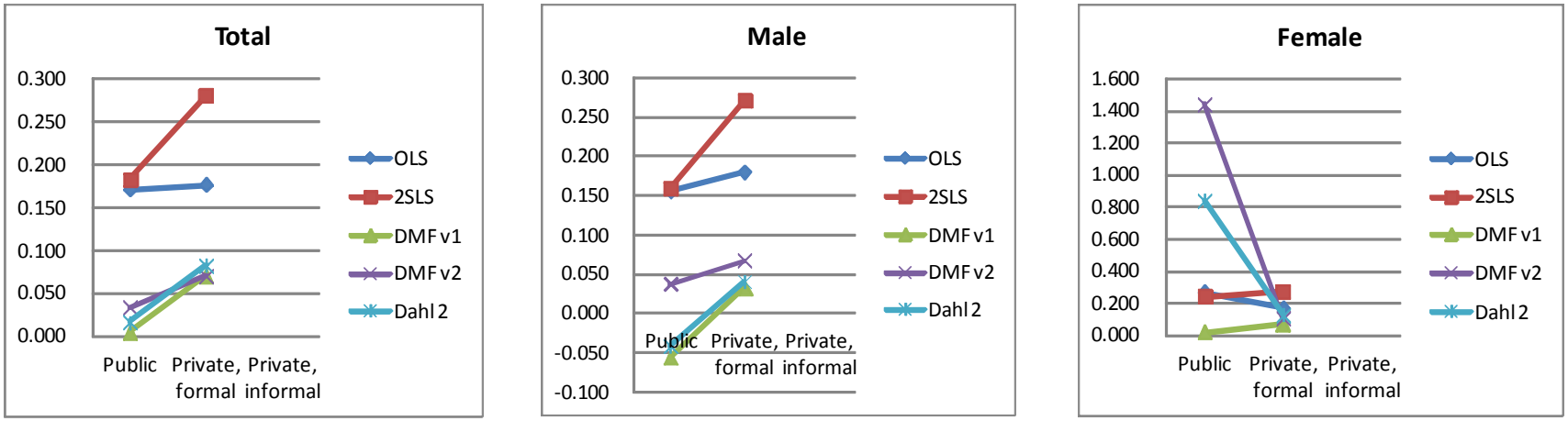
Notes: The sample used the models are only married workers. Not enough sample size is available for the informal sector for Asians.

Figure 8.5. The Marginal Return to Schooling in the Public, Formal Private and Informal Private in Different Models 1, Asians

(OLS, 2SLS, and Lee, DMF, and Dahl 1)



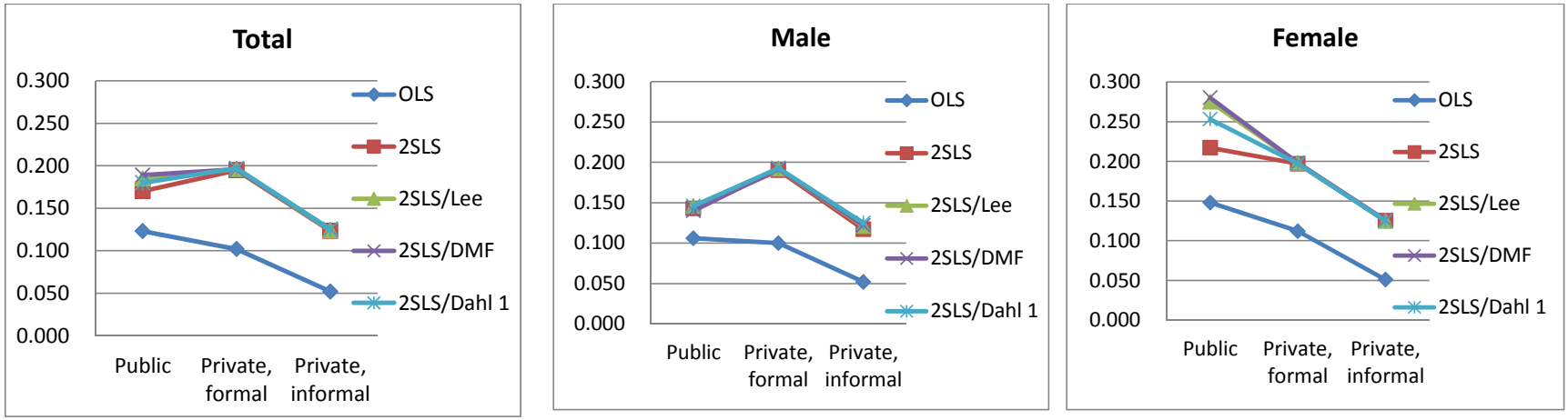
(OLS, 2SLS, and DMF variant 1, DMF variant 2, and Dahl 2)



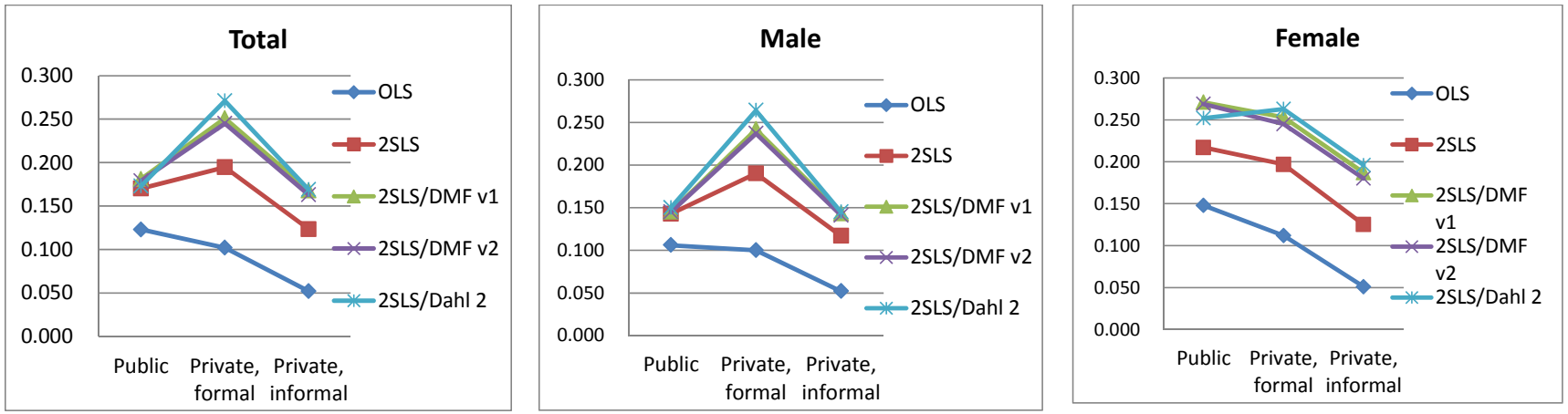
Notes: The models include only married workers. Not enough sample size is available for the informal sector for Asians.

Figure 8.6. The Marginal Return to Schooling in the Public, Formal Private and Informal Private in Different Models 2, All Races

(OLS, 2SLS, and 2SLSs with Lee, DMF, and Dahl 1)



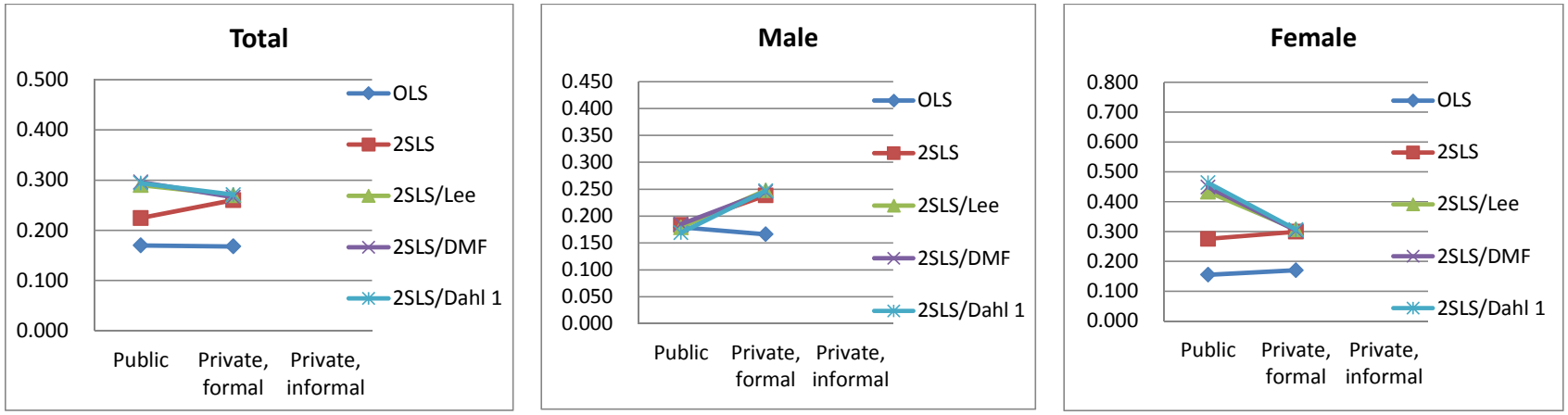
(OLS, 2SLS, and 2SLSs with DMF variant 1, DMF variant 2, and Dahl 2)



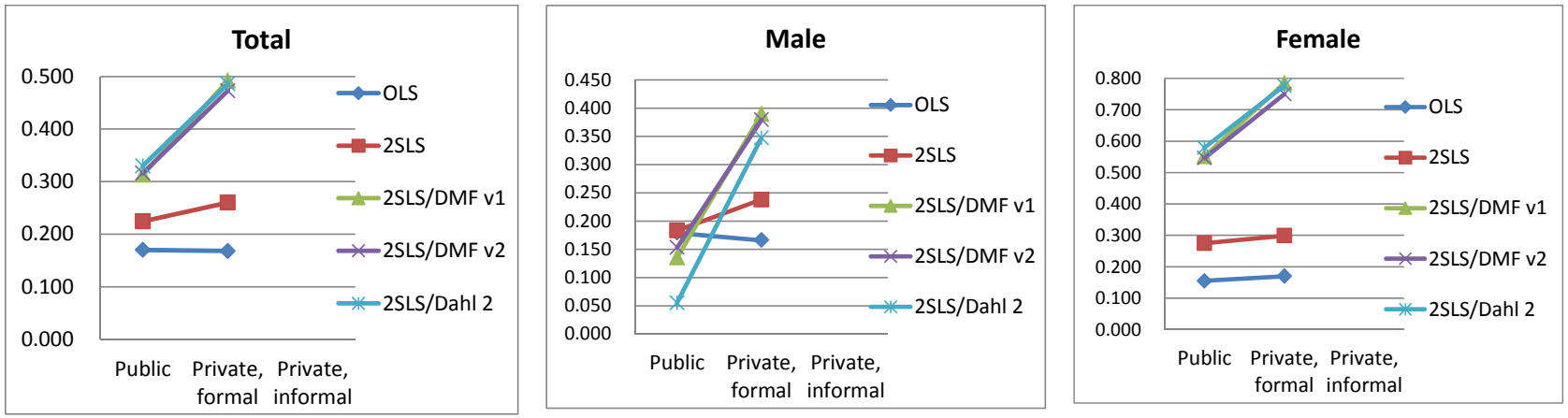
Note: The sample used the models are only married workers.

Figure 8.7. The Marginal Return to Schooling in the Public, Formal Private and Informal Private in Different Models 2, Whites

(OLS, 2SLS, and 2SLSs with Lee, DMF, and Dahl 1)



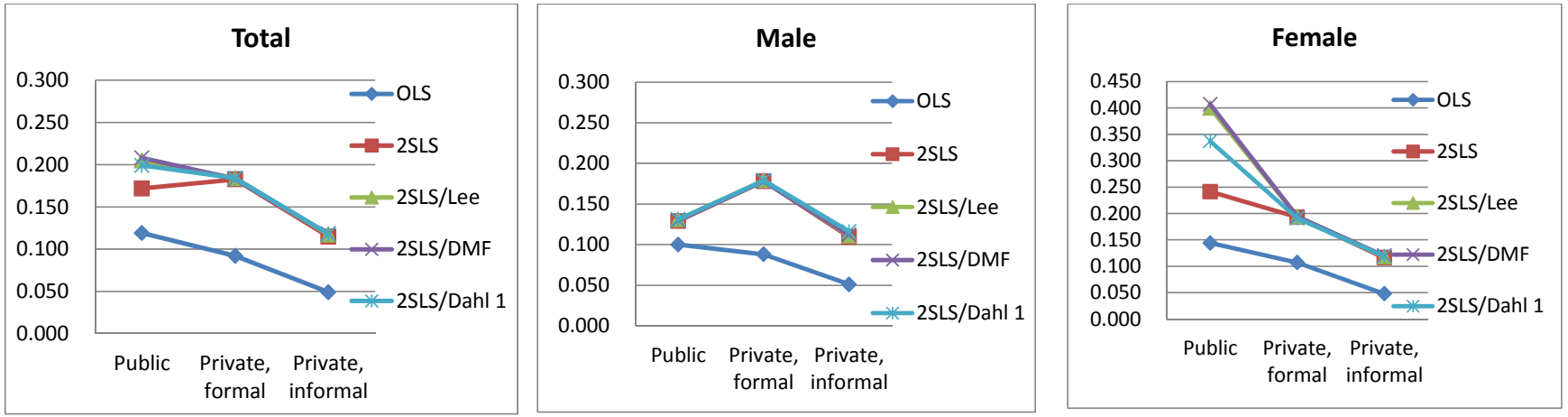
(OLS, 2SLS, and 2SLSs with DMF variant 1, DMF variant 2, and Dahl 2)



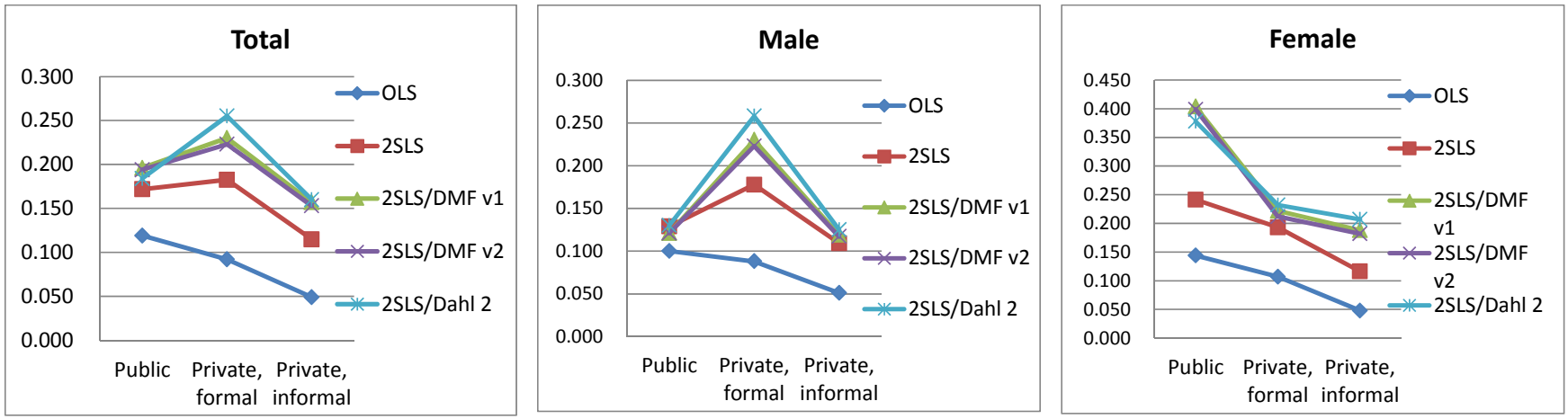
Note: The sample used the models are only married workers. Not enough sample size is available for the informal sector for Whites

Figure 8.8. The Marginal Return to Schooling in the Public, Formal Private and Informal Private in Different Models 2, Blacks

(OLS, 2SLS, and 2SLSs with Lee, DMF, and Dahl 1)



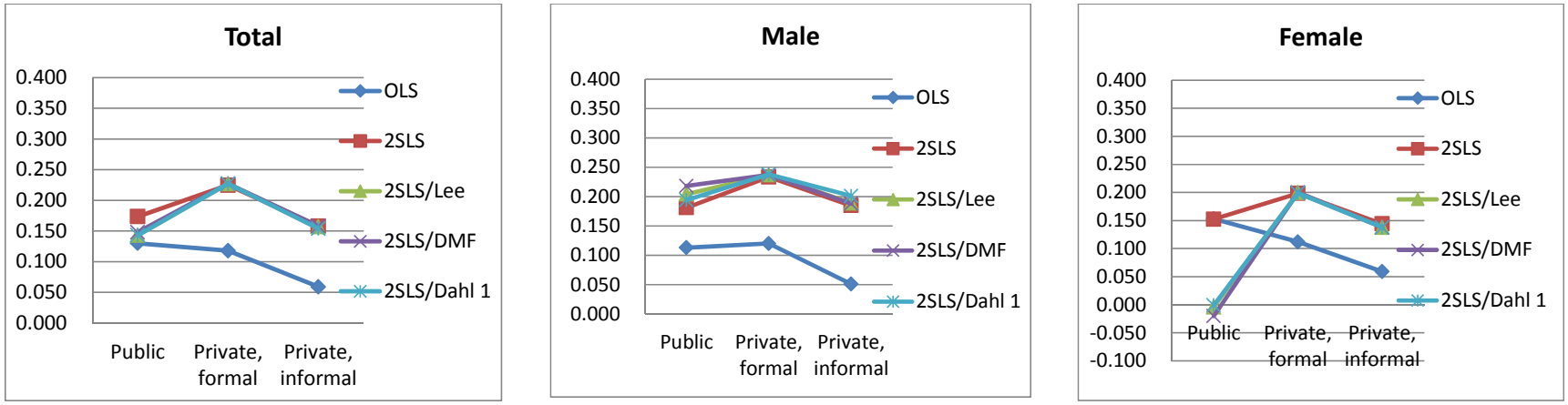
(OLS, 2SLS, and 2SLSs with DMF variant 1, DMF variant 2, and Dahl 2)



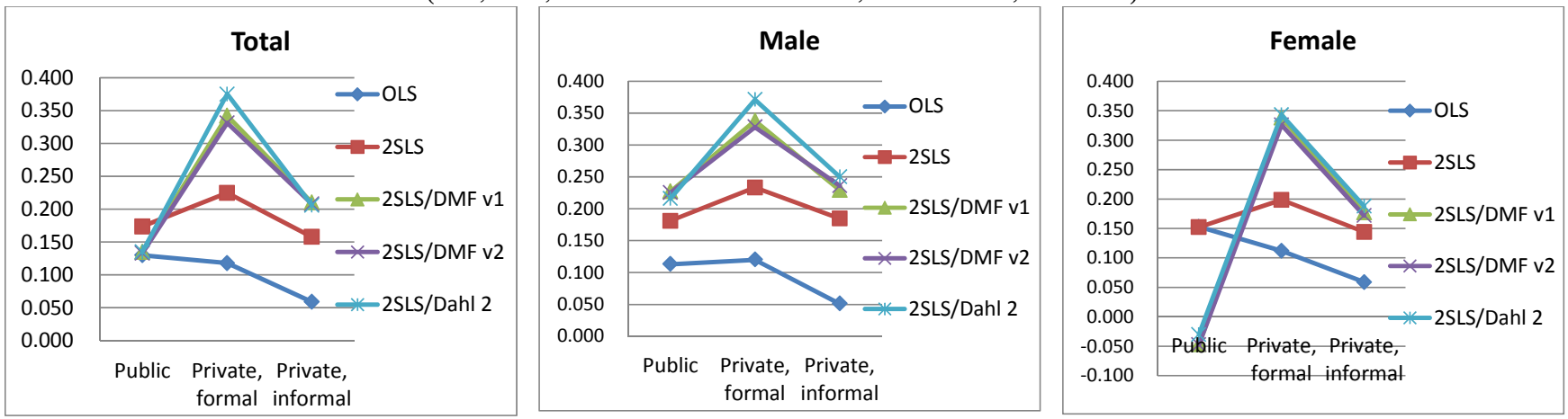
Note: The sample used the models are only married workers.

Figure 8.9. The Marginal Return to Schooling in the Public, Formal Private and Informal Private in Different Models 2, Coloreds

(OLS, 2SLS, and 2SLSs with Lee, DMF, and Dahl 1)



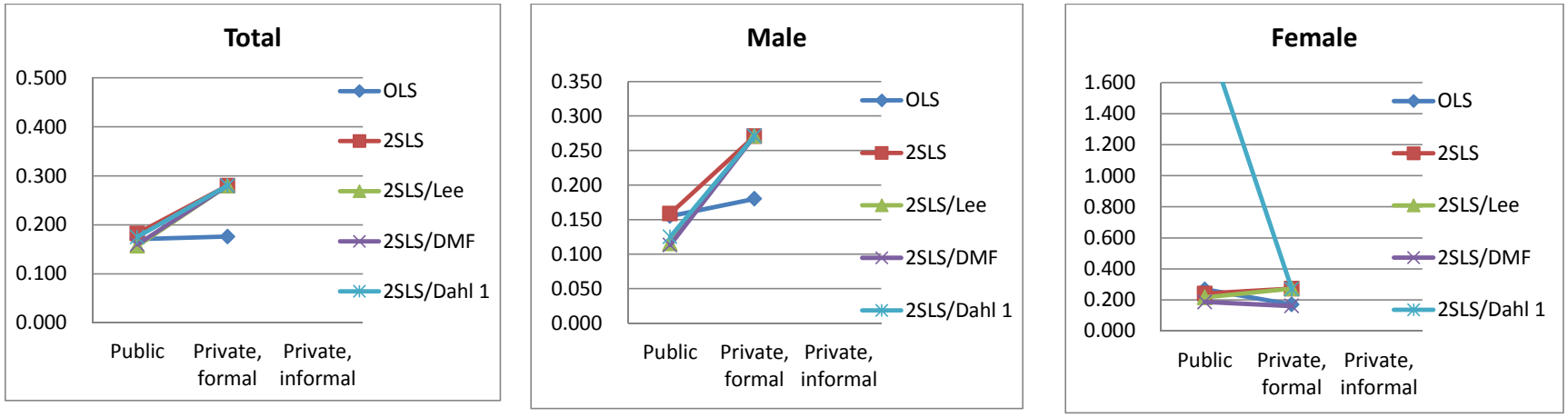
(OLS, 2SLS, and 2SLSs with DMF variant 1, DMF variant 2, and Dahl 2)



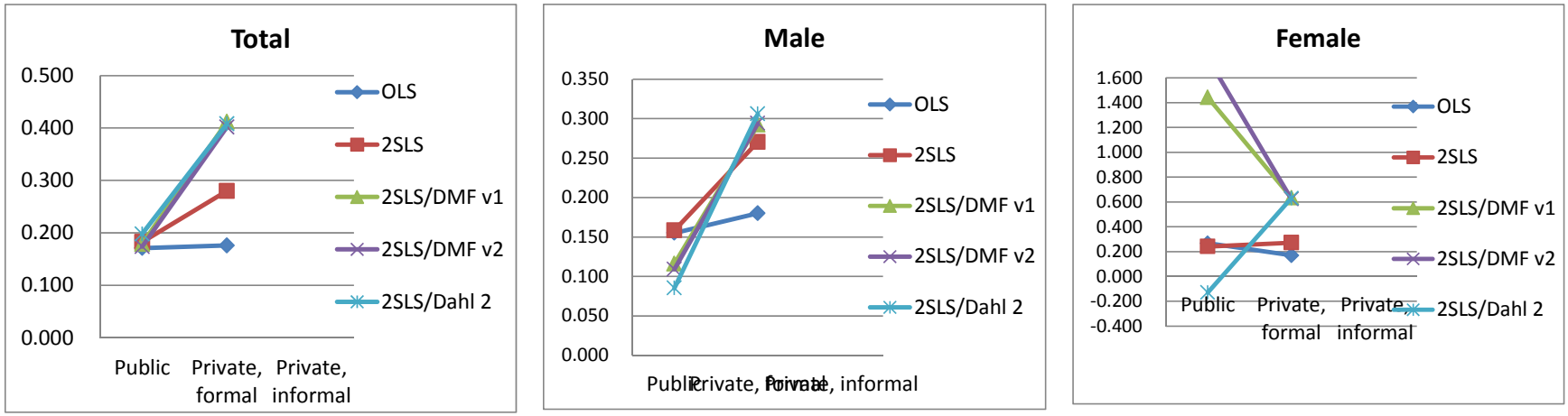
Notes: The sample used the models are only married workers. The estimated marginal return to schooling in the public sector for Colored female is not statistically significant.

Figure 8.10. The Marginal Return to Schooling in the Public, Formal Private and Informal Private in Different Models 2, Asians

(OLS, 2SLS, and 2SLSs with Lee, DMF, and Dahl 1)



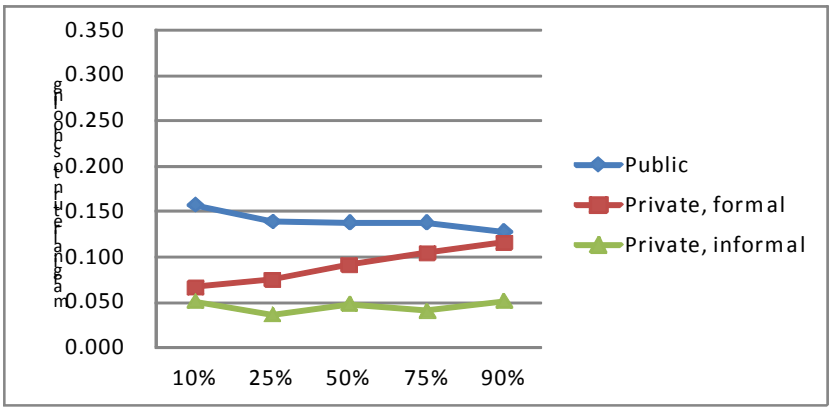
(OLS, 2SLS, and 2SLSs with DMF variant 1, DMF variant 2, and Dahl 2)



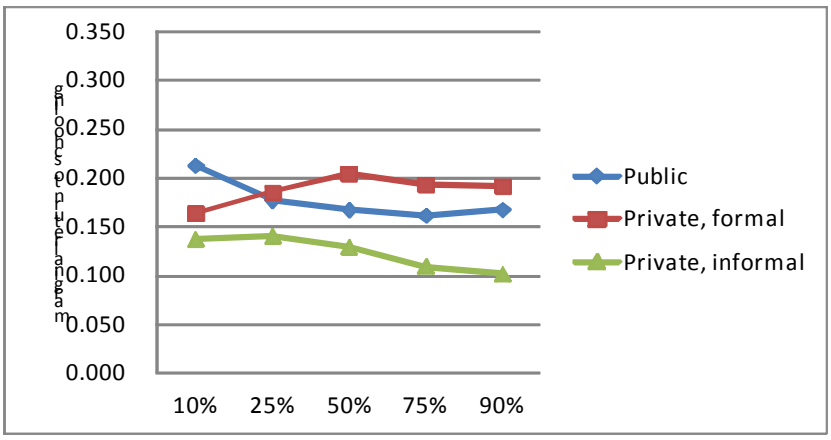
Notes: Only married workers are included. Coefficients which are not statistically significant are also included.

Figure 8.11. Marginal Returns to Schooling at Different Quantiles by Sector

(Quantile Regression)



(Instrumental Variable Quantile Regression)



(Instrumental Variable Quantile Regression with Selection Correction)

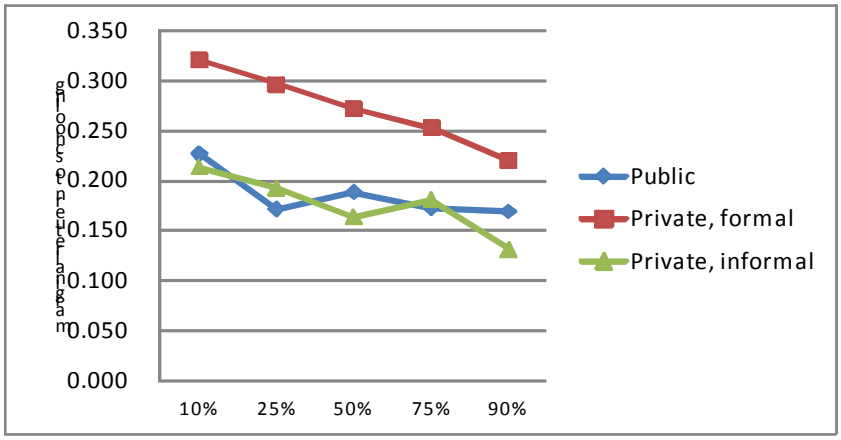
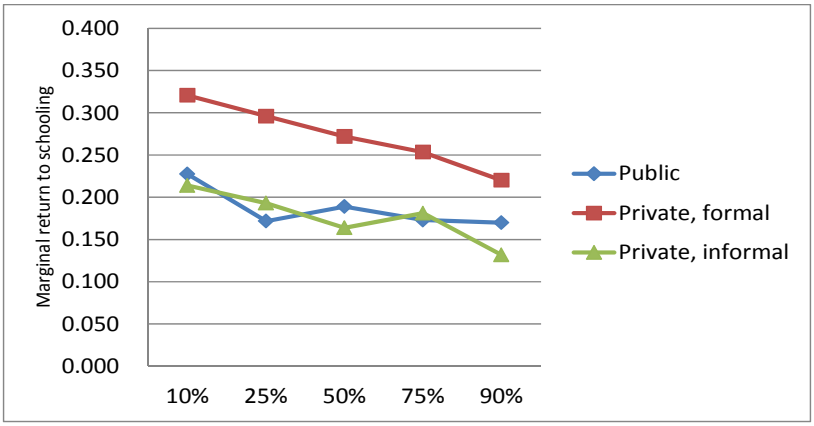
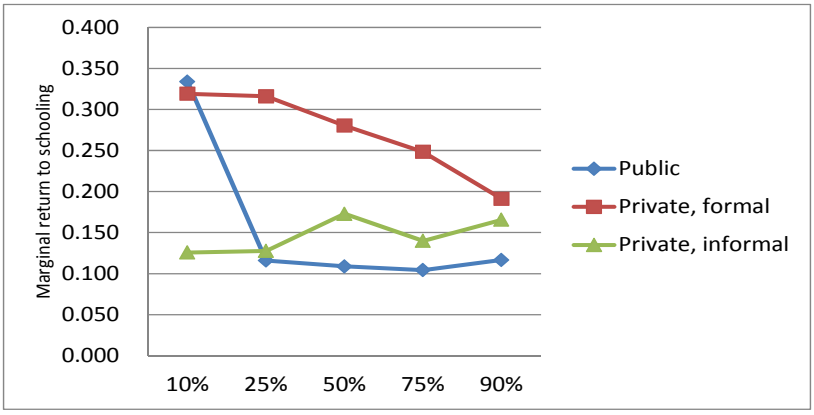


Figure 8.12. Marginal Return to Schooling at Different Quantiles by Sector and Gender, Using Instrumental Variable Quantile Regression with Selection Correction

(Total)



(Male)



(Female)

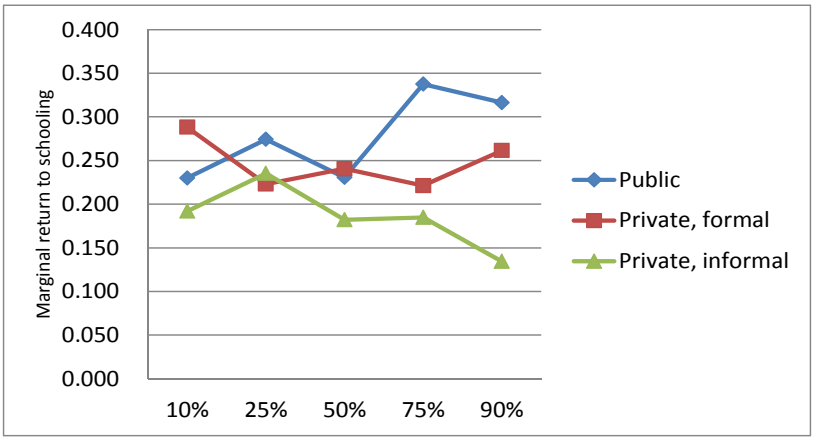
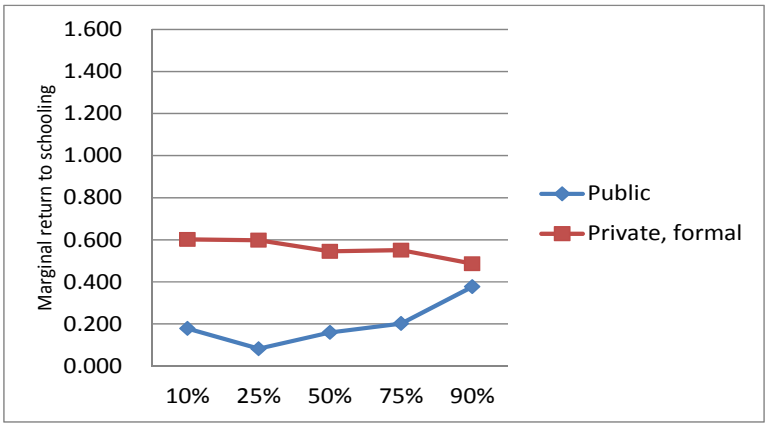
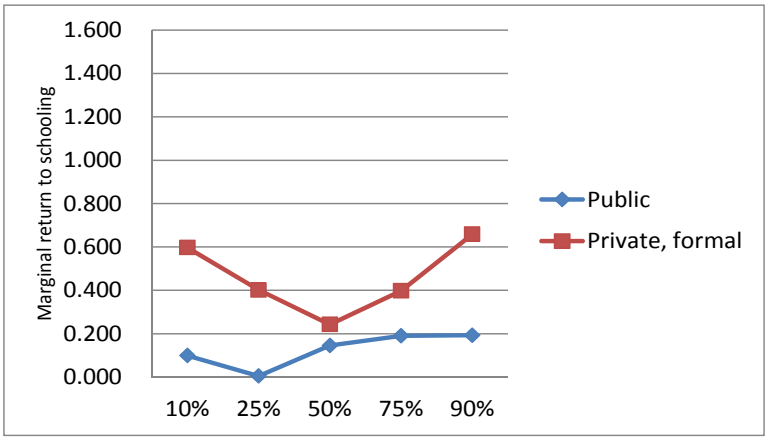


Figure 8.13.Marginal Return to Schooling at Different Quantiles by Sector for Whites, Using Instrumental Variable Quantile Regression with Selection Correction

(Total)



(Male)



(Female)

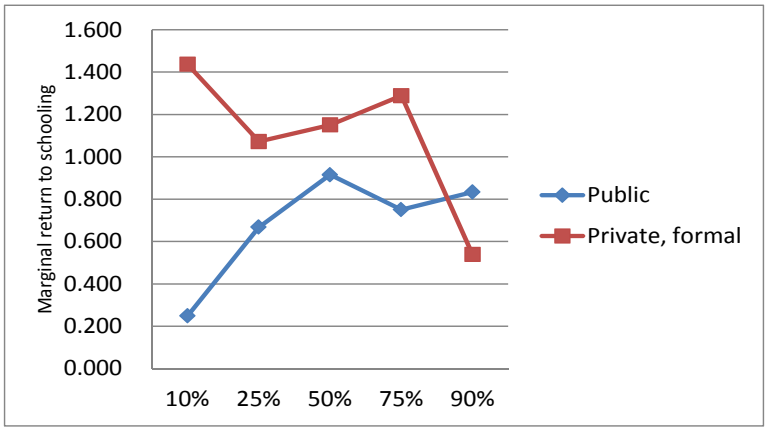
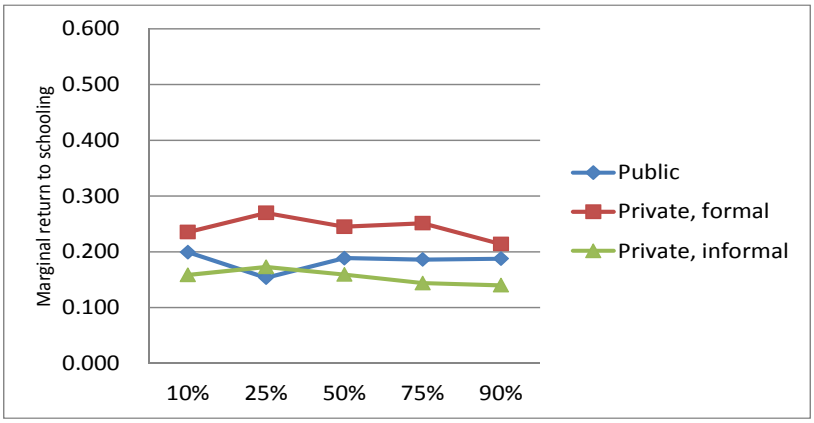
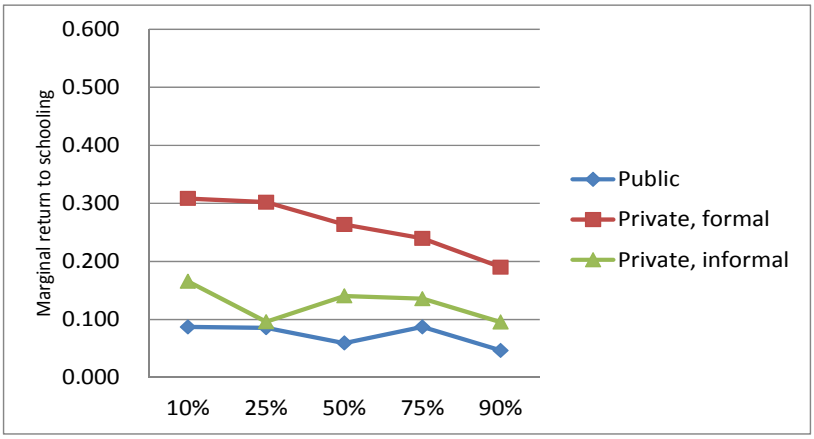


Figure 8.14. Marginal Return to Schooling at Different Quantiles by Sector for Blacks, Using Instrumental Variable Quantile Regression with Selection Correction

(Total)



(Male)



(Female)

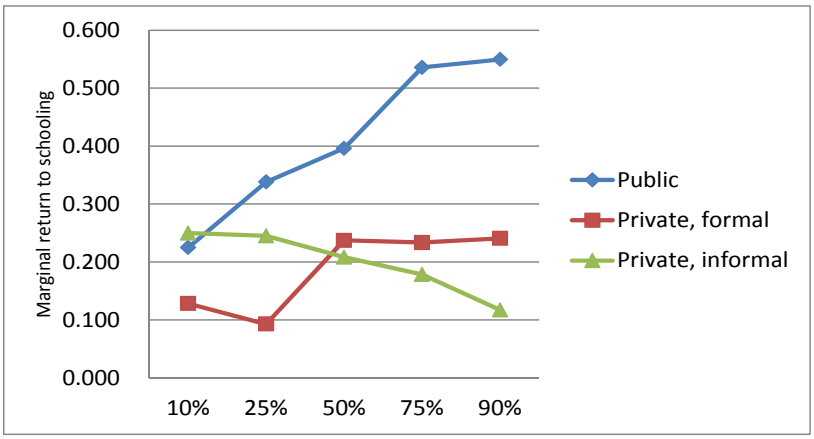
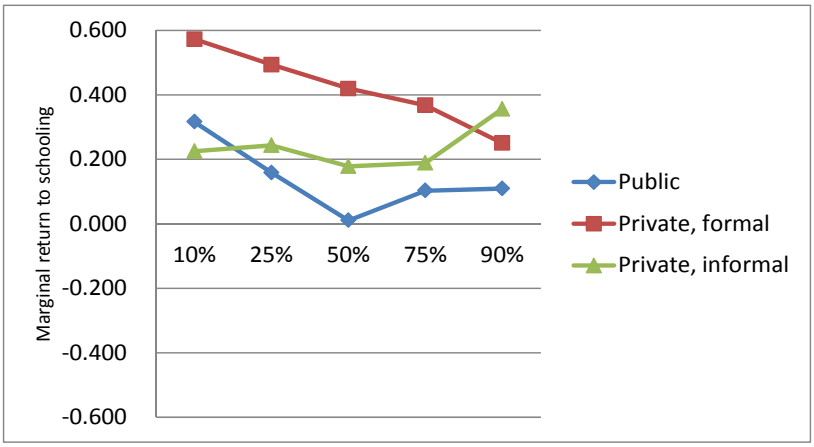
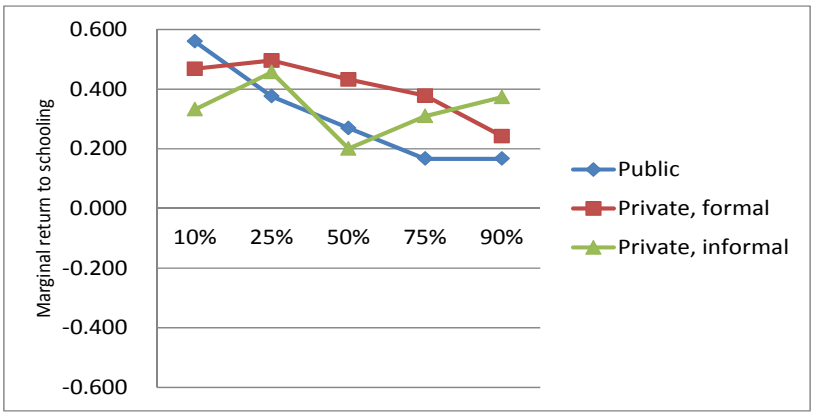


Figure 8.15. Marginal Return to Schooling at Different Quantiles by Sector for Colored, Using Instrumental Variable Quantile Regression with Selection Correction

(Total)



(Male)



(Female)

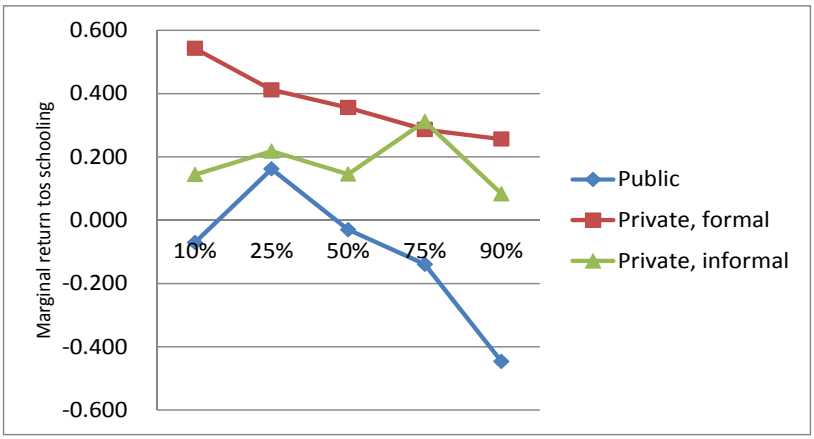
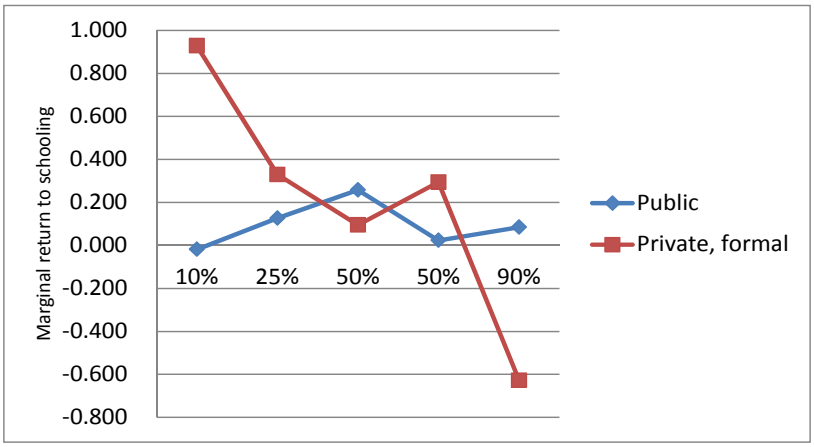
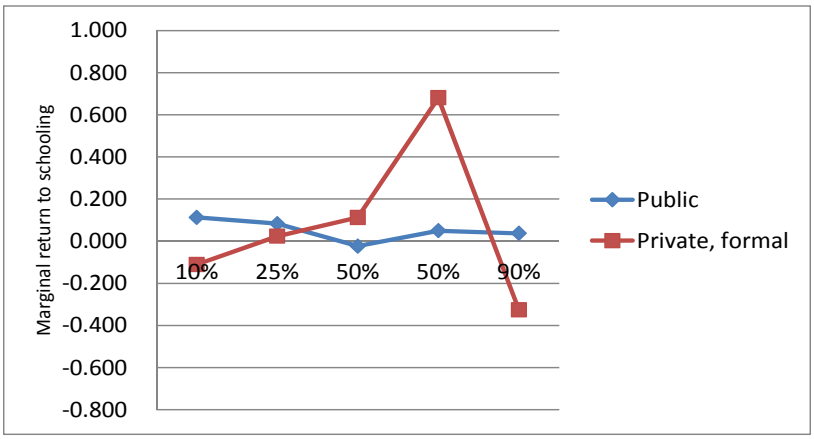


Figure 8.16. Marginal Return to Schooling at Different Quantiles by Sector for Asians, Using Instrumental Variable Quantile Regression with Selection Correction

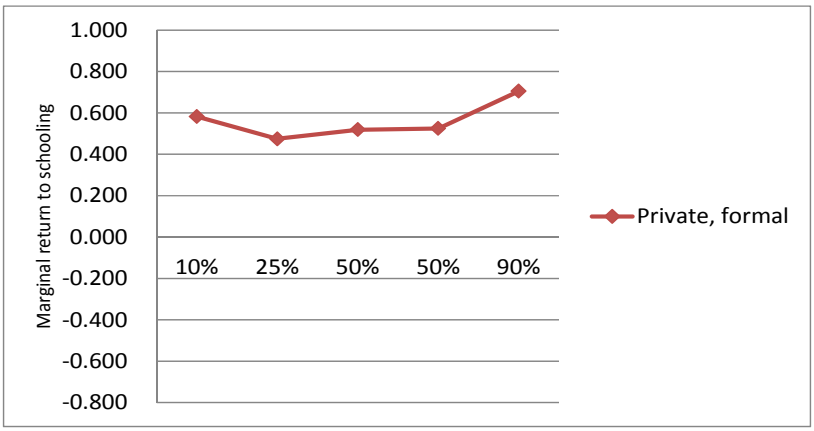
(Total)



(Male)



(Female)



9. RESULTS OF THE ESTIMATION OF THE EFFECT OF YEARS OF SCHOOLING ON INFORMAL SECTOR EMPLOYMENT

This chapter presents the results of the estimation of the effects of increased years of schooling on the probability of being employed in the public, formal private, and informal private sectors. The estimates of a multinomial logit choice model with four categories (public, formal private, informal private, and out of labor force/unemployed) based on equation 6.21 in Chapter 6 are shown first. Then, the results are presented for a probit choice model with different combinations of two sector categories (public vs. formal private, public vs. informal private, and formal private vs. informal private) based on equation 6.20 in Chapter 6. The probit estimation was conducted for the total sample and for married people separately. In addition, the instrumental variable probit model was conducted to control for endogeneity of education.

In general, greater schooling had a positive effect on the probability of employment in the public and formal private sectors and a negative effect on employment in the informal private sector relative to other sectors of employment. Also the analysis here suggests that educated people who cannot find a job in the public or formal sector possibly choose to be unemployed or out of labor force. More schooling substantially raises the probability of working in the formal private and public sectors relative to being unemployed/out of labor force, but not as much in the informal sector. Being more educated does not results in greater employment probability in the informal private sector, relative to being unemployed/out of labor force, especially for women.

9.1. Results of the Multinomial Logit Choice Model

Table 9.1 shows the estimates of the multinomial logit sector choice model with three

sector categories (public, formal private, and informal private) and out of labor force/ unemployed category. This model did not control for the endogeneity of schooling since the instrumental variable multinomial choice model does not have a well-developed method yet. The endogeneity of schooling was controlled in the next model, discussed in the next section. The analysis was conducted with the out of labor force/unemployed as a base category and the marginal effects are calculated based on the estimate. The marginal effects were usually statistically significant at the 0.1% level. The coefficients were statistically significant at the 0.1% level except for the ones in the informal private sector which is either statistically significant at the 5% level or insignificant.

The results show that workers with more schooling were more likely to work in the public and formal private sectors (relative to all other options)[40]. However, they also show that workers with greater schooling were less likely to be working in the informal private sector. With an additional year of schooling, workers were 0.5% points more likely to work in the public sector, 2.6% points more likely to work in the formal private sector, but 0.7% points less likely to work in the informal private sector relative to being unemployed/out of the labor force, and 2.5% points less likely to be out of labor force or unemployed.

In addition, estimates of the education coefficients in the multinomial logit choice model by race and gender are shown in Table 9.2. The coefficients were also statistically significant at either the 0.1%, 1% or 10% level, with a few exceptions. The results indicate that workers with more schooling were more likely to work in the public and formal private sectors when compared to being unemployed/out of the labor force, but less likely to work in the informal

---

[40] A marginal effect of schooling in a sector means the increase in the probability of being in the sector given additional year of schooling, compared to changes in probabilities of all other options.

sector relative to being unemployed/out of the labor force across race groups.

Tables 9.3 and 9.4 show the coefficients in the multinomial logit choice model for all races, by race and gender. For total sample, the coefficient of schooling is 0.41 for the public, 0.14 for the formal private, 0.01 for the informal private, which means that an additional schooling greatly increases the probabilities of working in the public and formal private sectors compared to being unemployed and out of labor force, while it slightly increases the probability of working in the informal private sector compared to being unemployed and out of labor force. The coefficient of schooling in the informal private sector is even smaller and statistically insignificant for female, Whites, and Asians. This suggests that a substantial number of people who could not find a job in the public or formal private sector possibly choose to be unemployed or out of labor force instead of working in the informal sector.

9.2. Results of the Probit Choice Model

Next, the estimates of the probit model are presented. The findings show that workers with more schooling were more likely to work in the public and formal private sectors relative to the informal private sector. This was true across gender and race, except for White males. Also the effect of schooling on employment in the formal private sector relative to the employment in the informal sector, was much stronger for females than males, except for Whites and Asians.

9.3. Results of the Instrumental Variable Probit Model

The instrumental variable probit model was used because education could also be endogenous in determining employment probabilities.

9.3.1. Test for the instrument

Table 9.3 shows the Wald statistics to test for the exogeneity of years of schooling in the estimation of the effect of schooling on the likelihood of working in the public sector versus the

formal private sector, on the likelihood of working in the public sector versus the informal private sector, and on the likelihood of working in the formal private sector versus the informal private sector. The Wald statistics tested the null hypothesis that years of schooling was an exogenous variable. If the null hypothesis is rejected, the instrumental variable selected earlier, spouse schooling, would be used.

Regarding the model comparing the public and formal private sectors, the Wald statistics shows that one can reject the null hypothesis that years of schooling is exogenous in the estimation involving the total sample, female sample, White male sample, and Black female sample at the 10% (or lower %) significance level. However, we cannot reject the null hypothesis for other sub-samples including the male sample.

The Wald statistics of the model comparing the public and informal private sectors shows that one can reject the null hypothesis that years of schooling is exogenous for the total sample, female sample, total Black sample, and total Colored sample at the 10% (or less %) level. However, one cannot reject the null hypothesis for the other samples.

In the estimation comparing the formal private and informal private sectors, one can reject the null hypothesis that years of schooling is exogenous for the total sample, the female sample, total Black sample, Black female sample, total Colored sample, and female Colored sample. However, one cannot reject the null hypothesis for the other samples.

9.3.2. Results

Table 9.4 shows the marginal effects of years of schooling on public sector employment relative to the formal sector employment estimated by probit and IV probit. Given the variation in exogeneity tests reported in the previous section, some of the samples should use just the probit model and not the IV probit model, but all the results are shown, for information purposes.

The IV probit coefficients are highlighted.

Table 9.5 reports the marginal effects of years of schooling on public sector employment relative to the informal sector employment. The results were very similar across different samples and also statistically significant in general (mostly at the 0.1% level). The overall finding is that more education increases the probability of working in the public sector versus working in the informal sector. The impact of education was larger for females across race groups and for Whites. In the IV probit model, an additional year of schooling increased the likelihood of working in the public sector versus working in the informal private sector by 7.1% points for total, 4.8% points for males, and 9.2% points for females.

The marginal effects of years of schooling on formal private sector employment relative to informal sector employment are shown in Table 9.6. The results shown are generally statistically significant. The marginal effects become larger in the IV probit compared to the ordinary probit for married people, but they were similar between the probit model for the total sample and the IV probit. Again, additional schooling increased the likelihood of working in the formal private sector, compared to working in the informal private sector. The effect of schooling was stronger for females. An additional year of schooling increased the probability of working in the formal private sector, compared to the informal private sector, by 1.5% points for total, 0.9% points for males, and 5.6% points for females in the IV probit model.

9.4. Robustness Check

To further confirm the robustness of the results on the probability of working in the formal private sector versus working in the informal private sector, several additional estimations were conducted. First, the definition of the informal sector was changed, as was also done in the robustness check of the returns to schooling findings in an earlier chapter. Second, language and

district dummies were added to the model. In sum, the coefficients are robust across different models shown below.

Table 9.7 shows marginal effects of years of schooling on the likelihood of working in the formal sector versus working in the informal sector estimated by probit and IV probit model using the workers in the informal sector and/or workers without written contract as the definition of the informal sector. The results confirm the findings in the previous model. It even increased the number of statistically significant marginal effects such as marginal effects for Black males in the IV probit and Colored total in the probit model. Also the marginal effects of years of schooling on the employment in the formal sector increased.

Table 9.8 presents the coefficients of the probit and IV probit sector choice model using the workers in the informal sector and/or without any of the written contract, medical aid/health insurance, and pension/retirement fund as the definition of the informal sector. Again the patterns of results are similar to those discussed earlier.

Table 9.9 presents the estimated coefficients of the model using the definition of informal employment as: workers without any of the written contract, medical aid/health insurance, and pension/retirement fund. Again, the findings are similar to previous models.

Table 9.10 displays the probit/IV probit estimations with district dummies. The coefficients are very similar to the original models.

9.5. Summary and Conclusion

This chapter investigated the effect of schooling on the likelihood of working in the informal sector compared to being in other options. It also examined the likelihood of being employed in the informal sector compared to one of the other two sectors of employment.

As expected, schooling had a negative effect on the probability of working in the

informal sector when compared to the formal and public sectors. These findings were consistent in the models with and without controlling for the endogeneity of education. Moreover, these findings were robust across the models with different set of variables and different definitions for informal sector

The dissertation used a multinomial logit model to estimate the likelihoods of working in the public, formal private, informal private sectors, and being unemployed/out of labor force. The results found that an additional year of schooling was associated with an increase in the likelihood of working in the public sector relative to all other options by 0.5%, an increase in the likelihood of working in the formal private sector by 2.6%, and a decrease in the likelihood of working in the informal private sector relative to all other options by 0.7% for the total sample. The trends were similar for the male and female samples in general. Also the coefficients in the analysis imply that a lot of educated people who cannot find a job in the public or formal sector possibly choose to be unemployed or out of labor force.

The results just discussed appear to suggest that schooling does not generally have a strong role in producing an improved likelihood of employment for workers in the informal sector, at least when compared to the private and public sectors. The fact that increased just slightly raises the probability of working in the informal sector relative to the probability of being unemployed or out of the labor force imply that more-educated workers prefer to wait for a higher-paying job in the formal private or public sectors (remaining unemployed or just waiting outside the labor force) instead of taking a job in the informal sector. This suggests a relatively weak informal sector in terms of providing rewarding job opportunities for skilled workers. The last chapter will provide a set of recommendations regarding this result.

Table 9.1 *Estimations of the Multinomial Logit Sector Choice Model (Public, Formal Private, and Informal Private), Marginal Effects, All Races*

All Races	Total				Male				Female			
	Public	Formal private	Informal private	Out of labor force/ unemployed	Public	Formal private	Informal private	Out of labor force/ unemployed	Public	Formal private	Informal private	Out of labor force/ unemployed
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
years of schooling	0.0054***	0.026***	-0.0067***	-0.025***	0.0047***	0.023***	-0.0079***	-0.020***	0.0058***	0.040***	-0.0053***	-0.041***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)
N	63785				29453				34332			

Notes: * p<0.05, ** p<0.01, *** p<0.001.Standard errors are in the parentheses.

Table 9.2 *Estimations of the Multinomial Logit Sector Choice Model (Public, Formal Private, and Informal Private), Marginal Effects, by Race*

Whites	Total				Male				Female			
	Public	Formal private	Informal private	Out of labor force/ unemployed	Public	Formal private	Informal private	Out of labor force/ unemployed	Public	Formal private	Informal private	Out of labor force/ unemployed
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
years of schooling	0.016***	0.029***	-0.014**	-0.032***	0.0090**	0.042***	-0.016*	-0.035***	0.032***	0.032***	-0.002	-0.062***
	(0.003)	(0.005)	(0.004)	(0.004)	(0.003)	(0.007)	(0.006)	(0.008)	(0.006)	(0.008)	(0.002)	(0.008)
N	3875				1903				1972			
Blacks	Total				Male				Female			
	Public	Formal private	Informal private	Out of labor force/ unemployed	Public	Formal private	Informal private	Out of labor force/ unemployed	Public	Formal private	Informal private	Out of labor force/ unemployed
	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)
years of schooling	0.0048***	0.025***	-0.0090***	-0.021***	0.0054**	0.017***	-0.0081*	-0.014***	0.002	0.021***	-0.0046*	-0.019***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.001)	(0.005)	(0.002)	(0.004)
N	49745				22768				26977			

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parentheses.

Table 9.2 (continued)

Colored	Total				Male				Female			
	Public	Formal private	Informal private	Out of labor force/ unemployed	Public	Formal private	Informal private	Out of labor force/ unemployed	Public	Formal private	Informal private	Out of labor force/ unemployed
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
years of schooling	0.015***	0.026***	-0.0040**	-0.037***	0.0099***	0.017***	0.001	-0.028***	-0.048***	0.016***	0.040***	-0.0082**
	(0.003)	(0.003)	(0.001)	(0.003)	(0.003)	(0.004)	(0.001)	(0.004)	(0.005)	(0.004)	(0.004)	(0.003)
N	8804				4091				4713			
Asian	Total				Male				Female			
	Public	Formal private	Informal private	Out of labor force/ unemployed	Public	Formal private	Informal private	Out of labor force/ unemployed	Public	Formal private	Informal private	Out of labor force/ unemployed
	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)
years of schooling	0.024***	0.065***	-0.016*	-0.072***	0.024**	0.055***	-0.008	-0.071***	0.007	0.092***	-0.010	-0.090***
	(0.007)	(0.010)	(0.008)	(0.010)	(0.008)	(0.012)	(0.008)	(0.013)	(0.005)	(0.015)	(0.011)	(0.015)
N	1361				691				670			

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parentheses.

Table 9.3 *Estimations of the Multinomial Logit Sector Choice Model (Public, Formal Private, and Informal Private), Coefficients, All Races*

All Races									
	Total			Male			Female		
	Public	Formal private	Informal private	Public	Formal private	Informal private	Public	Formal private	Informal private
	(1)	(2)	(3)	(5)	(6)	(7)	(9)	(10)	(11)
years of schooling	0.410***	0.142***	0.011*	0.307***	0.110***	0.016*	0.524***	0.176***	0.002
	(0.007)	(0.004)	(0.004)	(0.010)	(0.006)	(0.007)	(0.012)	(0.007)	(0.006)
N		63785			29453			34332	

Notes: * p<0.05, ** p<0.01, *** p<0.001.Standard errors are in the parentheses.

Table 9.4 *Estimations of the Multinomial Logit Sector Choice Model (Public, Formal Private, and Informal Private), Coefficients, by Race*

Whites									
	Total			Male			Female		
	Public	Formal private	Informal private	Public	Formal private	Informal private	Public	Formal private	Informal private
	(1)	(2)	(3)	(5)	(6)	(7)	(9)	(10)	(11)
years of schooling	0.324***	0.237***	0.042	0.222***	0.200***	-0.015	0.436***	0.265***	0.111
	(0.030)	(0.022)	(0.042)	(0.044)	(0.033)	(0.050)	(0.043)	(0.031)	(0.079)
N		3875			1903			1972	
Blacks									
	Total			Male			Female		
	Public	Formal private	Informal private	Public	Formal private	Informal private	Public	Formal private	Informal private
	(13)	(14)	(15)	(17)	(18)	(19)	(21)	(22)	(23)
years of schooling	0.444***	0.135***	0.011*	0.333***	0.108***	0.012	0.566***	0.164***	0.007
	(0.009)	(0.005)	(0.005)	(0.012)	(0.007)	(0.008)	(0.014)	(0.008)	(0.006)
N		49745			22768			26977	

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parentheses.

Table 9.4 (continued)

Colored									
	Total			Male			Female		
	Public	Formal private	Informal private	Public	Formal private	Informal private	Public	Formal private	Informal private
	(1)	(2)	(3)	(5)	(6)	(7)	(9)	(10)	(11)
years of schooling	0.406***	0.160***	0.059***	0.325***	0.104***	0.087***	0.482***	0.209***	0.018
	(0.018)	(0.010)	(0.014)	(0.024)	(0.014)	(0.021)	(0.029)	(0.016)	(0.018)
N		8804			4091			4713	
Asian									
	Total			Male			Female		
	Public	Formal private	Informal private	Public	Formal private	Informal private	Public	Formal private	Informal private
	(13)	(14)	(15)	(17)	(18)	(19)	(21)	(22)	(23)
years of schooling	0.659***	0.367***	0.117	0.537***	0.306***	0.093	0.832***	0.408***	0.067
	(0.064)	(0.038)	(0.067)	(0.081)	(0.054)	(0.082)	(0.114)	(0.060)	(0.129)
N		1361			691			670	

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parentheses.

Table 9.5 *Test of Exogeneity of Schooling*

All Races	Public vs Formal Private			Public vs Informal Private			Formal vs Informal Private		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Wald test of exogeneity	7.83	1.23	16.82	12.17	1.89	6.38	36.56	0.99	60.63
	p=0.0051	p=0.2666	p=0.0000	p= 0.0005	p=0.1691	p=0.0115	p=0.0000	p=0.3195	p=0.0000
N	8474	5636	2838	4852	2546	2306	8983	5607	3376
Whites	Public vs Formal Private			Public vs Informal Private			Formal vs Informal Private		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Wald test of exogeneity	0.66	3.42	0.00	0.21	0.74	0.05	0.13	0.08	0.37
	p=0.4150	p=0.0644	p=0.9528	p=0.6500	p= 0.3889	p= 0.8184	p= 0.7208	p=0.7717	p=0.5456
N	1773	1033	740	420	217	203	1517	912	605
Blacks	Public vs Formal Private			Public vs Informal Private			Formal vs Informal Private		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
Wald test of exogeneity	0.00	0.52	4.31	6.40	1.71	2.27	8.84	0.09	18.12
	0.961	0.470	0.038	p= 0.0114	p=0.1914	p=0.1317	p=0.0029	p=0.7624	p=0.0000
N	4437	3116	1321	3499	1818	1681	3473	1895	1569
Colored	Public vs Formal Private			Public vs Informal Private			Formal vs Informal Private		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)
Wald test of exogeneity	1.07	0.71	0.00	3.54	1.24	0.72	3.40	0.46	11.42
	p=0.3002	p=0.3994	p=0.9635	p=0.0598	p=0.2664	p=0.3946	p=0.0651	p= 0.4968	p= 0.0007
N	1835	1194	641	830	-	394	1815	-	717
Asians	Public vs Formal Private			Public vs Informal Private			Formal vs Informal Private		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)
Wald test of exogeneity	1.86	1.28	0.32	0.09	0.26	-	0.33	0.10	0.22
	p=0.1721	p= 0.2573	p=0.5741	p=0.7670	p=0.6093	-	p=0.5686	p=0.7521	p=0.6367
N	429	1194	136	98	71	-	408	285	102

Table 9.6. *Estimations of Probit and IV Probit Sector Choice Model, Marginal Effects, Public vs. Formal Private*

All Races	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.013***	0.010***	0.041***	0.014***	0.009***	0.054***	0.014***	0.009***	0.038***
	(0.001)	(0.001)	(0.004)	(0.001)	(0.001)	(0.005)	(0.001)	(0.001)	(0.005)
N	17786	10645	7106	9774	6478	3296	8474	5636	2838
Whites	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.013***	0.003	0.035***	0.011**	0.002	0.037***	0.007	-0.010	0.036*
	(0.004)	(0.005)	(0.008)	(0.004)	(0.005)	(0.009)	(0.007)	(0.008)	(0.016)
N	2386	1349	1036	1813	1051	760	1773	1033	740
Blacks	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.018***	0.017**	0.034**	0.018**	0.018***	0.021	0.028***	0.021***	0.057***
	(0.005)	(0.005)	(0.013)	(0.005)	(0.005)	(0.016)	(0.003)	(0.004)	(0.008)
N	11324	6901	4397	5601	3893	1708	4437	3116	1321
Colored	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.025***	0.020***	0.032***	0.025***	0.013***	0.052***	0.032***	0.023***	0.050***
	(0.004)	(0.004)	(0.006)	(0.004)	(0.004)	(0.008)	(0.006)	(0.006)	(0.013)
N	3440	1976	1456	1905	1219	685	1835	1194	641
Asians	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.039***	0.036**	0.037*	0.044***	0.035**	0.120**	0.017	0.023***	0.085
	(0.010)	(0.013)	(0.017)	(0.011)	(0.011)	(0.039)	(0.017)	(0.006)	(0.060)
N	612	396	209	441	298	137	429	1194	136

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parenthesis. Only the IV probit estimations in highlighted cells are better than ordinary probit estimates.

Table 9.7 *Estimations of the Provit and IV Probit Sector Choice Model, Marginal Effects, Public vs. Informal Private*

All Races	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.074***	0.049***	0.094***	0.067***	0.049***	0.095***	0.071***	0.048***	0.092***
	(0.002)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.005)	(0.006)
N	11272	4999	6233	5829	2876	2953	4852	2546	2306
Whites	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.057***	0.074***	0.026*	0.049***	0.060**	0.045*	0.047*	0.044	0.058*
	(0.012)	(0.017)	(0.012)	(0.013)	(0.019)	(0.020)	(0.019)	(0.028)	(0.029)
N	570	296	274	423	216	207	420	217	203
Blacks	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.047***	0.032***	0.046*	0.051***	0.044***	0.019	0.064***	0.042***	0.062***
	(0.010)	(0.010)	(0.020)	(0.011)	(0.007)	(0.019)	(0.004)	(0.005)	(0.006)
N	9095	3911	5147	4415	2130	2285	3499	1818	1681
Colored	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.065***	0.049***	0.082***	0.059***	0.045***	0.082***	0.071***	0.052***	0.077***
	(0.005)	(0.007)	(0.007)	(0.008)	(0.008)	(0.010)	(0.009)	-0.011	(0.014)
N	1438	675	753	871	444	423	830	436	394
Asians	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.12***	0.092**	0.50*	0.14***	0.12***	0.14	0.13*	0.10	-
	(0.027)	(0.028)	(0.200)	(0.031)	(0.036)	(2007.7)	(0.058)	(0.067)	-
N	151	101	40	104	74	20	98	71	-

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parenthesis. Only the IV probit estimations in highlighted cells are better than ordinary probit estimates.

Table 9.8. *Estimations of the Provit and IV Probit Sector Choice Model, Marginal Effects, Formal Private vs. Informal Private*

All Races	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.013***	0.015***	0.021***	0.0072***	0.0075***	0.029***	0.015***	0.0086***	0.056***
	(0.001)	(0.002)	(0.003)	(0.001)	(0.002)	(0.004)	(0.002)	(0.002)	(0.006)
N	20428	11491	8937	10520	6466	4054	8983	5607	3376
Whites	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.016***	0.030***	0.0036	0.010**	0.016**	0.0031	0.014*	0.021*	0.011
	(0.005)	(0.008)	(0.003)	(0.004)	(0.006)	(0.007)	(0.006)	(0.009)	(0.014)
N	2054	1209	845	1545	928	617	1517	912	605
Blacks	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.030***	0.023***	0.038***	0.022***	0.016***	0.035***	0.012***	0.0015	0.036***
	(0.001)	(0.002)	(0.003)	(0.002)	(0.003)	(0.006)	(0.003)	(0.003)	(0.006)
N	14291	7971	6320	6647	4114	2533	3473	1895	1569
Colored	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.012***	0.0015	0.036***	0.0094**	0.0014	0.038***	0.021***	0.00013	0.072***
	(0.003)	(0.003)	(0.006)	(0.003)	(0.003)	(0.007)	(0.005)	(0.025)	(0.011)
N	3473	1895	1569	1895	1116	762	1815	1096	717
Asians	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.039**	0.038**	0.037	0.041***	0.041**	0.064	0.041***	0.041**	0.031
	(0.012)	(0.014)	(0.021)	(0.012)	(0.013)	(0.038)	(0.012)	(0.013)	(0.081)
N	578	389	165	408	285	102	408	285	102

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parenthesis. Only the IV probit estimations in highlighted cells are better than ordinary probit estimates.

Table 9.9. *Estimations of the Probit and IV Probit Sector Choice Model, Marginal Effects, Formal Private vs. Informal Private, Using the Workers in the Informal Sector and/or Workers without Written Contract as the Definition of the Informal Sector*

All Races	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.028***	0.026***	0.041***	0.020***	0.016***	0.045***	0.035***	0.026***	0.065***
	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)	(0.005)
N	20428	11491	8937	10520	6466	4054	8983	5607	3376
Whites	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.021***	0.029***	0.016*	0.018***	0.025***	0.016	0.022*	0.034**	0.019
	(0.005)	(0.008)	(0.007)	(0.005)	(0.007)	(0.011)	(0.009)	(0.012)	(0.019)
N	2062	1209	849	1553	928	621	1524	912	609
Blacks	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.034***	0.027***	0.036***	0.029***	0.023***	0.037***	0.029***	0.023***	0.045***
	(0.002)	(0.002)	(0.005)	(0.002)	(0.002)	(0.007)	(0.002)	(0.002)	(0.005)
N	14291	7971	6320	6647	4114	2533	6647	4114	1929
Colored	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.027***	0.017***	0.048***	0.021***	0.008	0.050***	0.038***	0.04	0.075***
	(0.003)	(0.004)	(0.006)	(0.004)	(0.004)	(0.006)	(0.006)	(0.022)	(0.012)
N	3481	1906	1569	1895	1122	762	1815	1098	717
Asians	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.042***	0.041**	0.047*	0.035**	0.034*	0.065	0.031	0.024	0.079
	(0.012)	(0.014)	(0.023)	(0.014)	(0.015)	(0.044)	(0.027)	(0.029)	(0.074)
N	578	389	188	408	285	121	391	270	121

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parenthesis.

Table 9.10. *Estimations of the Probit and IV Probit Sector Choice Model, Marginal Effects, Formal Private vs. Informal Private, Using the Workers in the Informal Sector and/or without Any of the Written Contract, Medical Aid/Health Insurance, and Pension/Retirement Fund as the Definition of the Informal Sector*

All Races	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.025***	0.025***	0.038***	0.016***	0.014***	0.045***	0.033***	0.025***	0.069***
	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)	(0.005)
N	20428	11491	8937	10520	6466	4054	8983	5607	3376
Whites	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.022***	0.038***	0.012*	0.016***	0.022***	0.013	0.020**	0.031**	0.019
	(0.005)	(0.009)	(0.006)	(0.004)	(0.006)	(0.009)	(0.007)	(0.010)	(0.017)
N	2062	1209	849	1553	928	621	1524	912	609
Blacks	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.036***	0.029***	0.038***	0.031***	0.025***	0.038***	0.031***	0.025***	0.047***
	(0.002)	(0.002)	(0.005)	(0.002)	(0.002)	(0.007)	(0.002)	(0.002)	(0.005)
N	14291	7971	6320	6647	4114	2533	6647	4114	1929
Colored	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.027***	0.017***	0.049***	0.021***	0.009*	0.053***	0.040***	0.044*	0.084***
	(0.003)	(0.004)	(0.006)	(0.004)	(0.004)	(0.007)	(0.006)	(0.022)	(0.012)
N	3477	1903	1569	1895	1122	762	1815	1098	717
Asians	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.048***	0.040**	0.080**	0.041**	0.032*	0.137*	0.036	0.029	0.129
	(0.012)	(0.015)	(0.029)	(0.013)	(0.014)	(0.060)	(0.025)	(0.025)	(0.106)
N	578	389	188	408	285	121	391	270	121

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parenthesis.

Table 9.11. *Estimations of the Probit and IV Probit Sector Choice Model, Marginal Effects, Formal Private vs. Informal Private, Using "Without Any of the Written Contract, Medical Aid/Health Insurance, and Pension/Retirement Fund" as the Definition of the Informal Employment*

All Races	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.017***	0.017***	0.022***	0.012***	0.009***	0.029***	0.028***	0.026***	0.045***
	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.004)	(0.004)	(0.005)	(0.006)
N	16034	9302	6732	8061	5159	2902	6962	4422	2540
Whites	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.011**	0.018*	0.013*	0.011*	0.015*	0.011	0.009	0.019*	0.004
	(0.004)	(0.009)	(0.007)	(0.004)	(0.007)	(0.013)	(0.005)	(0.008)	(0.021)
N	1604	909	692	1168	681	484	1146	670	474
Blacks	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.032***	0.033***	0.033***	0.024***	0.027***	0.027*	0.024***	0.027***	0.043***
	(0.002)	(0.002)	(0.003)	(0.005)	(0.005)	(0.011)	(0.005)	(0.005)	(0.007)
N	10735	6338	4397	4828	3231	1597	4828	3231	1299
Colored	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.022***	0.017***	0.032***	0.019***	0.009*	0.042***	0.038***	0.061**	0.067***
	(0.004)	(0.004)	(0.006)	(0.004)	(0.004)	(0.007)	(0.007)	(0.024)	(0.012)
N	3217	1742	1470	1732	1021	706	1657	995	662
Asians	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.035**	0.027	0.073*	0.02	0.012	0.136**	0.026	0.019	-
	(0.013)	(0.018)	(0.030)	(0.011)	(0.013)	(0.052)	(0.019)	(0.025)	-
N	462	217	164	316	135	88	308	124	-

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parenthesis.

Table 9.12. *Estimations of the Probit and IV Probit Sector Choice Model, Marginal Effects, Formal Private vs. Informal Private, With District Dummies*

All Races	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.012***	0.012***	0.023***	0.008***	0.007**	0.033***	0.016***	0.009**	0.058***
	(0.002)	(0.002)	(0.005)	(0.002)	(0.002)	(0.007)	(0.004)	(0.003)	(0.009)
N	20428	11491	8937	10520	6466	4054	8983	5607	3376
Whites	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.018	0.036*	0.000	0.014	0.027	0.000	-	-	-
	(0.012)	(0.018)	0.000	(0.012)	(0.023)	0.000	-	-	-
N	1884	1030	578	1260	596	393	-	-	-
Blacks	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.027***	0.020***	0.032***	0.021***	0.014***	0.031**	-	-	-
	(0.002)	(0.002)	(0.005)	(0.003)	(0.003)	(0.011)	-	-	-
N	14291	7971	6320	6647	4104	2533	-	-	-
Colored	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.008	0.001	0.026	0.008	0.000	0.034*	-	-	-
	(0.004)	(0.002)	(0.015)	(0.005)	(0.003)	(0.015)	-	-	-
N	3408	1854	1506	1830	1074	722	-	-	-
Asians	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
years of schooling	0.040	0.000	0.058	0.000	0.063*	0.064	-	-	-
	(0.024)	0.000	(0.041)	(0.002)	(0.030)	(0.092)	-	-	-
N	484	308	126	321	218	66	-	-	-

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parenthesis.

10. GENERAL CONCLUSIONS, QUALIFICATIONS, AND POLICY RECOMMENDATIONS

This dissertation has examined the impact of education on earnings and employment in the public and private formal and informal sectors in South Africa, using the Labor Force Survey of September, 2007. The key research questions in this dissertation were: (1) To what extent does education affect the earnings of workers in the informal sector and how does the effect of education on earnings differ between the informal and formal private sector workers? and (2) To what extent does education affect the probability of formal and informal sector employment?

The returns to schooling were estimated using a variety of econometric methods intended to take into account possible endogeneity bias in the analysis of education, sample selection bias in the allocation of workers among various sectors of employment, including out of labor force/unemployed, and heterogeneity in rates of return to education. The most comprehensive model was the two-stage least squares model with multinomial two-step selection corrections to control for both endogeneity of schooling and sector sample selection bias. The quantile regression and piecewise linear spline function were applied to deal with the heterogeneity of returns to schooling. Moreover, the Blinder-Oaxaca decomposition method was conducted to explore the contribution of returns to schooling and educational endowments to wage differentials between sectors. To analyze the effects of schooling on the different sectors of employment, multinomial logit and probit models were estimated. The instrumental variable probit model was also used to control for the endogeneity of schooling.

In short, even after controlling for both endogeneity of education and sectoral sample selection bias, the return to schooling in the formal private sector was found to be much higher than that in the informal private sector. In fact, the rate of return to schooling was generally found to be the lowest in the informal private sector across race and gender groups. After controlling for the endogeneity of schooling and both sectoral and labor force participation sample selection bias using the 2SLS method with Durbin and Mcfadden selection correction, the marginal return to schooling in the informal sector was found to be equal to 16.8% for the overall sample studied, 14.3% for males and 18.7% for females, while that in the public sector was estimated to be equal to 18.1% for the total sample, 14.5% for males and 27.1% for females and that in the formal private sector was 25.1% for the total sample, 24.2% for males, and 25.35% for females. Also, the marginal return to schooling in the informal sector was slightly lower than the average for Black males, while being higher than the average for the Colored population.

Although not directly related to the focus of this dissertation, which is the comparative analysis of the formal and informal sectors, there was an interesting pattern observed in the rates of return to education in the public and formal private sectors, when disaggregated by gender. For males, the marginal return to schooling in the public sector was lower than that in the formal private sector, and this result remains true across race groups. By contrast, for females, the marginal return to schooling in the public sector was higher than that in the formal private sector.

The estimation using quantile regressions shows that conditional distributions of earnings given years of schooling were significantly different between the formal private and informal private sectors and do not often overlap in the model, even after controlling for both endogeneity

of schooling and sample selection bias. The conditional distributions of earnings given years of schooling were similar between the public and informal private sector.

The estimates of the piecewise spline function suggest convex marginal returns to education: schooling has a greater proportional impact on earnings as the level of education increases. This is especially true for the formal private and informal private sectors. The marginal returns to primary, lower secondary, upper secondary, and higher education in the informal private sector were 2.8%, 7.9%, 9.2%, and 40.8%, respectively. It is interesting to see that although rates of return to education were generally found to be lower in the informal sector, there is a significant exception: the marginal returns to higher education were the highest in the informal sector.

The results in the Blinder-Oaxaca decomposition show that the most significant factor in explaining wage differentials between the formal and informal sectors was not the differences in characteristics between the formal and informal sector, specially the higher levels of schooling of workers in the formal sector. Instead, it was the difference in the rewards offered to education and other characteristics of workers in the formal and informal sectors that explained the wage gaps between them. Education, for instance, is much more highly rewarded in the formal sector than in the informal sector, as was just discussed. The implications of this result are discussed below. At the same time, the dissertation also found that differences in average schooling tended to contribute much more to explaining wage differentials between the public and informal private sectors.

The study also found that increased schooling had a very small effect on the probability of working in the informal sector–relative to being unemployed or out of the labor force. By

contrast, the effect of increased schooling on working in the formal private and public sector –relative to unemployment/out of the labor force or relative to working in the informal sector– was positive. Also, the findings were consistent in the models with and without controlling for the endogeneity of education. These findings were robust across the models with different set of variables and different definitions for the informal sector. The interpretation of this result is examined below.

10.1. Interpretation and Policy Recommendations

The results just discussed appear to suggest that schooling does not generally have a strong role in producing improved labor market outcomes for workers in the informal sector, at least when compared to the private and public sectors. Therefore, putting everything together, the analysis supports the dual or segmented labor market view of the informal sector.

The fact that increased schooling does not substantially affect the probability of working in the informal sector relative to being unemployed or out of the labor force, implies that more-educated workers prefer to wait for a higher-paying job in the formal private or public sectors (remaining unemployed or just waiting outside the labor force) instead of taking a job in the informal sector. This suggests a relatively weak informal sector in terms of providing rewarding job opportunities for skilled workers.

The result showing lower rates of return to education in the informal sector, when compared to other sectors, also seems to indicate that employment in the informal sector consists of marginal, unskilled jobs where increased schooling does not produce greater earnings and income. With a lack of technological dynamism, the sector just fails to produce higher wages for

more-schooled workers.

It is likely that the workers that enter the informal sector (even those with higher levels of schooling) do so because they do not have access to the formal and public sectors, which are sheltered in a number of ways. One of the forces preventing access to the formal or public sectors may be linked to race. Blacks and Coloreds workers are much more likely to work in the informal private sector. This may be connected to racial discrimination in employment, geography and location, or differences in the quality of schooling. Another factor limiting employment in the formal sector may be trade unions, which are a powerful force in the formal private sector in South Africa. Many point out that the formal private sector labor market is distorted because of strong trade unions.

The government has expressed concern with the size of the informal sector and the fact that it is associated with poverty and low income. In terms of policy implications, there are two sets of possible policies that could be used to correct the difficulties faced by workers in the informal sector. First, the analysis on rates of return to education could be taken to imply that informal sector firms have relatively low productivity and lack technological dynamism. Insofar as moving from the informal to the formal sector may allow informal sector firms to upgrade their technology and to have access to public infrastructure and to public programs supporting private sector development, such a shift may raise wages and growth in the informal sector and in the overall South African economy. Therefore, efforts to provide incentives to informal sector firms to become formal may pay off in the form of greater productivity, for both the firms themselves as well as the government. The government could remove excessively complicated regulations on the labor market and business that lead some firms to move to the informal sector. Policymakers

should consider drastically reforming labor market policies and regulations in South Africa.

There is one set of results in the dissertation that do suggest that some of the employers in the informal sector may be avoiding the costs of onerous regulatory controls in the formal sector (including the penalties imposed by public sector corruption). This is the high rate of return that was observed in the informal sector for workers with a higher education degree, which exceeded that in the formal sector as well as in the public sector. This suggests that highly-skilled professionals (doctors, lawyers, etc.) may be operating out of the informal sector in order to avoid the difficulties and costs of dealing with the government bureaucracy.

Another policy implication is that the government should provide training and non-formal schooling for workers in the informal sector to complement the low quality of education they may have attained so that they will be become more productive. Assuming that improving the quality of schooling received by low-income families is a very long-run policy change, a more immediate solution for improving the earnings of informal workers is to provide them with greater training opportunities. Many of the workers in this sector have very low levels of schooling and their literacy skills may be very basic.

10.2. Qualifications and Implications for Future Studies

Conducting this research revealed some concerns about data and methodologies. First, it is possible that workers underreport their earnings and this could be more prevalent in a certain sector. Kuepie et al. (2009) suggest that measurement errors might appear in reported earnings, especially for non-salaried workers in the informal sector. Second, there could be measurement errors in reported years of schooling as well. Third, it is possible that some workers with specific characteristics do not report their earnings at all. Fourth, the taxes paid by workers were not

included in the survey questionnaire. Although the dissertation carried out some analysis examining the role of taxes in explaining differences in pay across sectors, future analysis should consider this issue with data that provide before-tax and after-tax information. Fourth, the validity of the instrument used to estimate the return to schooling is not completely confirmed. Future studies should analyze the validity of the instrument as well as look for better identification strategies. Finally, although it is out of the scope of this study, the issue of wages and employment in the public sector could be analyzed in greater detail in future studies.

REFERENCES

Amemiya, T. (1985). *Advanced econometrics*. Cambridge, MA: Harvard University Press.

Angrist, J. D., & Krueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings? *The Quarterly Journal of Economics, 106*(4), 979-1014.

Arias, O., & Khamis, M. (2008). *Comparative advantage, segmentation and informal earnings: A marginal treatment effects approach* (IZA Discussion Paper No. 3916). Bonn, Germany: Institute for the Study of Labor (IZA).

Ashenfelter, O., & Krueger, A. B. (1994). Estimates of the economic returns to schooling from a new sample of twins. *American Economic Review, 84*(5), 1157-1173.

Ashenfelter, O., & Rouse, C. (1998). Income, schooling, and ability: Evidence from a new sample of identical twins. *The Quarterly Journal of Economics, 113*(1), 253-284.

Ashenfelter, O., & Zimmerman, D. (1997). Estimating of return to schooling from sibling data: Fathers, sons and brothers. *Review of Economics and Statistics, 79*(1), 1-9.

Banerjee, B. (1983). The role of the informal sector in the migration process: A test of probabilistic migration models and labor market segmentation for India. *Oxford Economic Papers, New Series, 35*(3), 399-342.

Bargain, O., & Kwenda, P. (2009). *The informal sector wage gap: New evidence using quantile quartile? estimations on panel data* (Working Papers 200916). Dublin, Ireland: Geary Institute, University College Dublin.

Barro, R. J, & Lee, J-W. (2010). A new data set of educational attainment in the world, 1950-2010, NBER Working Paper Series 15902. Cambridge, MA: National Bureau of Economic Research.

Becker, G. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy, 70*(5), 9-49.

Becker, G. S. (1964). *Human capital.* New York, NY: Columbia University Press.

Blackburn, M., & Neumark, D. (1995). Are OLS estimates of the return to schooling biased downward? Another look. *Review of Economics and Statistics, 77*(2), 217-229.

Blinder, A. (1973). Wage discrimination: Reduced form and structural estimate. *Journal of Human Resources, 8*(4), 436-455.

Bhorat H., Lundall P., & Rospabé, S. (2002). The South African Labour Market in Economic and Legislative Considerations (Employment Paper 2002/32). Paris, France: International

Labour Organisation (ILO).

Borjas, G. J., & Bronars, S. G. (1989). Consumer discrimination and self-employment. *Journal of Political Economy, 97*(3), 581-605.

Bound, J., Jaeger, D., & Baker, R. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association 90*(430), 443–450.

Bourguignon, F., Fournier, M., & Gurgand, M. (2007). Selection bias corrections based on the multinomial logit model: Monte Carlo comparisons. *Journal of Economic Surveys, 21*(1), 174-205.

Buchinsky, M. (1998). Recent advances in quantile regression models: A practical guideline for empirical research. *The Journal of Human Resources, 33*(1), 88-126.

Butcher F. K, & Rouse, C. E. (2001). Wage effects of unions and industrial councils in South Africa. *The World Bank Policy Research Working Paper*, 2520

Butler, R. (1982). Estimating wage discrimination in the labor market. *Journal of Human Resources, 17*(4), 606-621.

Card, D. (1995). Earnings, schooling and ability revisited. In Solomon Polachek (Ed.), *Research in labor economics* (vol. 14) (pp. 23-48). Greenwich, CT: JAI Press.

Card, D. (1999). The causal effect of education on earnings. In O. Ashenfelter & D. Card (Eds.), *Handbook of labor economics* (vol. 3). Amsterdam, The Netherlands: North Holland.

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica, 69*(5), 1127-1160.

Chamberlain, D., & van der Berg, S. (2002). Earnings functions, labour market discrimination and quality of education in South Africa (Working Papers 02/2002). Stellenbosch, South Africa: Stellenbosch University, Department of Economics.

Charmes, J. (1990). A critical review of concepts, definitions and studies in the informal sector. In D. Turnham, B. Salomé, & A. Schwarz (Eds.), *The informal sector revisited*. Paris, France: OECD.

Charmes, J. (2000). Size, trends and productivity of women's work in the informal sector. Paper presented at the Annual IAFFE Conference, Istanbul.

Chen, M. A. (2007). Rethinking the informal economy: Linkages with the formal economy and the formal regulatory environment (DESA Working Paper No. 46). New York: United Nations.

Chernozhukov, V. & Hansen, C. (2004). Instrumental Variable Quantile Regression.

Chesher, A., Rosen, A., & Smolinski, K. (2011). An instrumental variable model of multiple discrete choice. Retrieved from http://econ-www.mit.edu/files/6657.

Clark, N. L. & Worger, W. H. (2004). South Africa: The Rise and Fall of Apartheid. Harlow, UK: Pearson Education Limited.

Cole, W., & Fayissa, B. (1991). The urban subsistence labor force: Toward a policy-oriented and empirically accessible taxonomy. *World Development, 19*(7), 779-789.

Cotton, J. (1988). On the decomposition of wage differentials. *Review of Economics and Statistics, 70*(2), 236-243.

Dahl, G.. (2002). Mobility and the returns to education: Testing a Roy model with multiple markets. *Econometrica, 70*(6), 2367-2420.

Davies, R., & Thurlow, J. (2009). Formal-informal economy linkages and unemployment in South Africa. Human Sciences Research Council, Centre for Poverty and Employment for Growth: Pretoria, South Africa.

De Soto, H. (1989). *The other path: The invisible revolution in the third world.* New York: Harper & Row.

De Soto, H. (2000). *The mystery of capital: Why capitalism triumphs in the west and fails everywhere else*. London: Black Swan.

De Soto, H. (2001). Dead capital and the poor. *SAIS Review, 21*(1), 13-43.

De Soto, H., & Litan, R. (2001). Effective property rights and economic development: Next steps. *Borookings-Wharton Papers on Financlal Services*, 251-271.

Development Policy Research Unit (DPRU). (2008). *Minimum wages, employment and household poverty: Investigating the impact of sectoral determinations*. Pretoria, South Africa: South Africa Department of Labor Development Policy Research Unit.

Dickens, W. T., & Lang, K. (1985). A test of dual labor market theory. *The American Economic Review, 75*(4), 792-805.

Dinkeleman, T. & Ranchhod, V. (2010). Evidence on the impact of minimum wage laws in an informal sector: domestic workers in South Africa. A Southern Africa Labour and Development Research Unit Working Paper Number 44. Cape Town: SALDRU, University of Cape Town.

Dubin, J. A., & McFadden, D. (1984). An econometric analysis of residential electric appliance

holdings and consumption. *Econometrica, 52*(2), 345-362.

Ewoudou, J., & Vencatachellum, D. (2006, November). An empirical analysis of private rates of returns to education in Cameroon (Working Paper*).* Tunis-Belvedère, Tunisia: African Development Bank.

Ferber, M. A., & Green, C. A. (1982). Traditional or reverse sex discrimination? A case study of a large public university. *Industrial and Labor Relations Review 35*(4), 550-564.

Fields, G. S. (1975). Rural-urban migration, urban unemployment and underemployment, and job-search activity in LDC's. *Journal of Development Economics, 2*(2), 165-187.

Fields, G. (1990). Labour market modelling and the urban informal sector: Theory and evidence. In D. Turnham, B. Salomé, & A. Schwarz (Eds.), *The informal sector revisited*. Paris, France: OECD.

Fields, G. S. (2005). A guide to multisector labor markets models. Paper prepared for the World Bank Labor Market Conference, Washington, DC.

Fitzenberger, B. (2003). Gender wage differences across quantiles, accounting for sample selection. Mimeo, University of Mannheim, Mannheim, Germany.

Funkhouser, E. (1996). The urban informal sector in Central America: Household survey evidence. *World Development, 24*(11), 1737-1751.

Funkhouser, E. (1997a). Demand-side and supply-side explanations for barriers to labor market mobility in developing countries: The case of Guatemala. *Economic Development and Cultural Change, 45*(2), 341-366.

Funkhouser, E. (1997b). Mobility and labor market segmentation: The urban labor market in El Salvador. *Economic Development and Cultural Change, 46*(1), 123-153.

Garen, J. (1984). The returns to schooling: A selectivity bias approach with a continuous choice variable. *Econometrica, 52*(5), 1199-1218.

Gindling, T. H. (1991). Labor market segmentation and the determination of wages in the public, private-formal., and informal sectors in San Jose, Costa Rica. *Economic Development and Cultural Change*, *39*(3), 585-605.

Goldin, C. and L. Katz (2007) *The Race between Education and Technology*. Cambridge, Massachusetts: Harvard University Press.

Gong, X., & van Soest, A. (2002). Wage differentials and mobility in the urban labour market: A panel data analysis for Mexico. *Labour Economics, 9*(4), 513-529.

Griffin, P., & Edwards, A. C. (1993). Rates of return to education in Brazil: Do labor market conditions matter? *Economics of Education Review, 12*(3), 245-256.

Griliches, Z. (1977). Estimating the returns to schooling: Some econometric problems. *Econometrica,* 45(1), 1-22.

Griliches, Z. (1979). Sibling models and data in economics: Beginnings of a survey. *Journal of Political Economy, 87*(5), 37-65.

Harris, J. R., & Todaro, M. P. (1970). Migration, unemployment, and development: A two-sector analysis. *American Economic Review, 60*(1), 126-142.

Hart, K. (1971). Informal income opportunities and urban employment in Ghana. In R. Jolly, E. de Kadt, H. Singer, & F. Wilson (Eds.), *Third world employment: Problems and strategy* (pp. 66-70). Harmondsworth, England: Penguin.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica 46*(6), 1251–1271.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica, 47*(1), 153-161.

Heckman, J., Ichimura, H., & Todd, P. (1998). Matching as an econometric estimator. *Review of Economic Studies, 65*(2), 261-94.

Heckman, J. J., & Vytlacil, E. J. (2001). Local instrumental variables. In C. Hsiao, K. Morimune, & J. L. Powell (Eds.), *Nonlinear statistical modeling: Proceedings of the Thirteenth International Symposium in Economic Theory and Econometrics: Essays in honor of Takeshi Amemiya* (pp. 1-46). New York, NY: Cambridge University Press.

Heckman, J. J., & Vytlacil, E. J. (2005). Structural equations, treatment effects and econometric policy evaluation. *Econometrica, 73*(3), 669-738.

Heckman, J. J., Urzua, S., & Vytlacil, E. J. (2006). Understanding instrumental variables in models with essential heterogeneity. *Review of Economics and Statistics, 88*(3), 389-432.

Idson, T., & Feaster, J. (1990). A selectivity model of employer-size wage differentials. *Journal of Labor Economics, 8*(1), 99-122.

Imbens, G. W. (2000). The role of the propensity score in estimating dose-response functions. *Biometrica, 87*(3), 706-710.

Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and Statistics, 86*(1), 4-29.

International Labor Organization (ILO). (1972). Employment, incomes and equality: A strategy

for increasing productive employment in Kenya. Geneva, Switzerland: ILO.

International Labor Organization. (1993). Resolution concerning statistics of employment in the informal sector, adopted by the Fifteenth International Conference of Labour Statisticians. Geneva, Switzerland: ILO.

Jann, B. (2008). The Blinder-Oaxaca decomposition for linear regression models. *The Stata Journal*, 8(4), 453-479.

Jones, P. (2001). Are educated workers really more productive? *Journal of Development Economics, 64*(1), 57-79.

Jütting, J., & Laiglesia, J. R. (2009). Employment, poverty reduction, and development: What's new? In J. Jütting & J. R. Laiglesia (Eds.), *Is informal normal? Towards more and better jobs in developing countries* (pp. 17-26). Paris, France: OECD.

Jütting, J. P., J. Parlevliet & T. Xenogiani (2008) Informal Employment Re-loaded" *OECD Development Center Working Papers*, No. 266.

Katz, L. and Autor, D. (1999) Changes in the Wage Structure and Wage Inequality. In O. Ashenfelter and D. Card (Eds.) *Handbook of Labor Economics*, Vol. 3A (pp. 1463-1555). Amsterdam, The Netherlands: North Holland.

Kingdon, G. & Knight, J. (2007). Unemployment in South Africa, 1995-2003: causes, problems, and policies. *Journal of African Economies* 16(5), 813-48.

Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica, 46*(1), 1-26.

Kruger, A. And Lindahl, M. (2001) Education for Growth: Why and For Whom?, *Journal of Economic Literature*, 39 (4), 1101-1136.

Kuepie, M., Nordman, C. J., & Roubaud, F. (2009). Education and earnings in urban West Africa. *Journal of Comparative Economics, 37*(3), 491-515.

Lam, D., & Schoeni, R.F. (1993). Effects of family background on earnings and returns to schooling: evidence from Brazil. *Journal of Political Economy, 101*(4), 710-740.

Lassibille, G., & Tan, J. (2005). The returns to education in Rwanda. *Journal of African Economies, 14*(1), 92-116.

Lechner, M. (2001). Identification and estimation of causal effects of multiple treatments under the conditional independence assumption. In M. Lechner &d F. Pfeiffer (Eds.), *Econometric evaluations of active labor market policies in Europe* (pp. 43-58). Heidelberg, Germany: Physical/Springer.

Lechner, M. (2002). Program heterogeneity and propensity score matching: An application to the evaluation of active labor market policies. *The Review of Economics and Statistics, 84*(2), 205-220.

Lee, L. (1983). Generalized econometric models with selectivity. *Econometrica, 51*(2), 507-512.

Leibbrandt, M., Woolard, I., & Woolard, C. (2007, January). Poverty and inequality dynamics in South Africa: post-apartheid developments in the light of the long-run legacy. Prepared for the IPC-DRCLAS workshop Brasilia 11-13 January 2007).

Leibbrandt, M., Woolard, I., McEwen, J., & Koep, C. (2010) Employment and inequality outcomes in South Africa. Cape Town: Southern Africa Labour and Development Research Unit (SALDRU) and School of Economics, University of Cape Town

Lewis, W. A. (1954). Economic development with unlimited supplies of labor. *Manchester School of Economic Studies, 22*(2), 139-191.

Loayza, N. V. (1996). The economics of the informal sector: A simple model and some empirical evidence from Latin America. *Carnegie-Rochester Conference Series on Public Policy, 45*(1), 129-162.

Maddala, G. S. (1983). *Limited-dependent and qualitative variables in econometrics*. Cambridge, England: Cambridge University Press.

Maloney, W. F. (2004). Informality revisited. *World Development, 32*(7), 1159-1178.

Mankiw, N. G., Romer, D., & Weil, D. N. (1992). A contribution of empirics of economic growth. *The Quarterly Journal of Economics, 107*(2), 407-437.

Marcoullier, D., Ruiz de Casilla, V., & Woodruff, C. (1997). Formal measures of the informal-sector wage gap in Mexico, El Salvador, and Peru. *Economic Development and Cultural Change, 45*(2), 367-392.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarenmbka (Ed.), *Frontiers in econometrics*. New York, NY: Academic Press.

Mincer, J. (1958). Investment in human capital and personal income distribution. *The Journal of Political Economy, 25*(1), 77-87.

Mincer, J. (1974). *Schooling, experience and earnings.* New York, NY: National Bureau of Economic Research and Columbia University Press.

Murphy, K. M., Shleifer, A., & Vishny, R. W. The allocation of talent: implication for growth. The Quarterly Journal of Economics, 106(2), 503-530.

Neuman, S., & Oaxaca R. (2004). Wage decompositions with selectivity-corrected wage equations: a methodological note. *Journal of Economic Inequality, 2*(1), 3-10.

Nakamuro, M. & Yamasaki, I. (2012). Return to education in the informal sector: an empirical study using South African Labor Force Survey. *Hikaku Kyouiku Gakkai Kenkyu* (*ComparativeEducation),* 44,88-108. (Written in Japanese).

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review, 14*(3), 693-709.

Oaxaca, R., & Ransom, M. (1999). Identification in detailed wage decomposition. *The Review of Economics and Statistics, 81*(1), 154-157.

O'Gorman, M. (2007). Educational disparity and racial earning iInequality. Toronto, Canada: University of Toronto.

Oosthuizen, M., & Bhorat, H. (2006). Educational outcomes in South Africa: A production function approach. Secretariat for Institutional Support for Economic Research in Africa (SISERA) Working Paper Series. Dakar, Senegal: SISERA.

Packard, T. G. (2007). Do workers in Chile choose informal employment? A dynamic analysis of sector choice (World Bank Policy Research Working Paper 4232). Washington, DC: World Bank.

Pagán, J. A., & Tijerina-Guajardo, J. (2000). Increasing wage dispersion and the changes in relative employment and wages in Mexico's urban informal sector: 1987-1993. *Applied Economics, 32*(3), 335-347.

Pencavel, J. (1998). Assortative mating by schooling and the work behavior of wives and husbands. *American Economic Review, 88*(2), 326-29.

Perry, G. E., Maloney, W. F., Arias, O. S., Fajnzylber, P., Mason, A. D., & Saavedra-Chanduvi, J. (2007). *Informality: Exit and exclusion.* Washington, DC: World Bank.

Pisani, M. J., & Pagán, J. A. (2004). Sectoral selection and informality: A Nicaraguan case study. *Review of Development Economics, 8*(4), 541-556.

Polgreen, L. (2012, March 23). In a deivided city, many blacks see echoes of white superiority. *New York Times International,* p. A6.

Rankin, N., Sandefur, J., & Teal, F. (2010). Learning and earning in Africa: Where are the returns to education high? CSAE Working Paper Series No. 2010-02. Oxford, England: University of Oxford.

Reimers, C. W. (1983). Labor market discrimination against Hispanic and black men. *The Review of Economics and Statistics, 65*(4), 570-579.

Rivera-Batiz, F.L. (2008). Education as an engine of economic development: Global experiences and prospects for El Salvador. FUSADES. San Salvador: El Salvador.

Rivers, D., & Vuong, Q. H. (1988). Limited information estimators and exogeneity tests for simultaneous probit models. *Journal of Econometrics 39* (3), 347-366.

Rosenbaum, P., & Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*, 41-55.

Roy, A. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers, 3*(2), 135-146.

Saavedra, J., & Chong, A. (1999). Structural reform, institutions and earnings: Evidence from the formal and informal sectors in urban Peru. *Journal of Development Studies, 35*(4), 95-116.

Schneider, F. (2006). *Shadow economies and corruption all over the world: What do we really know?* (IZA Discussion Paper No. 2315). Bonn, Germany: Institute for the Study of Labor (IZA).

Schultz, T. P. & Mwabu, G. (1998). Labor unions and the distribution of wages and employment in South Africa. *Industrial and Labor Relations Review, 51*(4), 680-703.

Seekings, J. & Nattras, N. (2006). *Class, Race, and Inequality in South Africa*. Scottsville, South Africa: University of KwaZulu-Natal Press.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

Shleifer, A. & Vishny, R. W. (1993). Corruption. *The Quarterly Journal of Economics, 108*(3), 599-617.

Smith, P. A., & Metzger, M. R. (1998). The return to education: Street vendors in Mexico. *World Development, 26*(2), 289-296.

Söderbom, M., Teal., F., Wambugu, A., & Kahyarara, G. (2006). Dynamics of returns to education in Kenyan and Tanzanian manufacturing. *Oxford Bulletin of Economics and Statistics, 68*(3), 261-288.

Solow, R. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics, 70*(1), 65-94.

South Africa History Online. Liberation struggle in South Africa. Retrieved from http://www.sahistory.org.za/liberation-struggle-south-africa/apartheid-and-limits-non-violent-resistance-1948-1960 (accessed on May 4, 2012)

Staiger, D., & Stock, J. H. (1997). Instrumental variables regression with weak instruments. E*conometrica, 65*(3), 557-586.

Statistics South Africa. (2003). Labor force survey March 2003. Pretoria, South Africa: Statistics South Africa.

Statistics South Africa. (2004a). Labor force survey September 2003. Pretoria, South Africa: Statistics South Africa.

Statistics South Africa. (2004b). Labor force survey March 2004. Pretoria, South Africa: Statistics South Africa.

Statistics South Africa. (2004c). Labor force survey September 2004. Pretoria¸ South Africa: Statistics South Africa.

Statistics South Africa. (2005a). Labor force survey March 2005. Pretoria, South Africa: Statistics South Africa.

Statistics South Africa. (2005b). Labor force survey September 2005. Pretoria, South Africa: Statistics South Africa.

Statistics South Africa. (2006a). Labor force survey March 2006. Pretoria, South Africa: Statistics South Africa.

Statistics South Africa. (2006b). Labor force survey September 2006. Pretoria, South Africa: Statistics South Africa.

Statistics South Africa. (2007a). Labor force survey March 2007. Pretoria, South Africa: Statistics South Africa.

Statistics South Africa. (2007b). Labor force survey: Round 16 September 2007 [Codebook SADA0152]. Pretoria, South Africa: Statistics South Africa.

Statistics South Africa. (2007c). Labor force survey September 2007 questionnaire. Pretoria, South Africa: Statistics South Africa.

Statistics South Africa. (2007d). Labor force survey September 2007. Pretoria, South Africa: Statistics South Africa.

Statistics South Africa. (2009). Historical revision September series 2000 to 2007. Pretoria, South Africa: Statistics South Africa.

Statistics South Africa. (2010). Quarterly labor force survey: quarter 1, 2010 (P0211). Pretoria, South Africa: Statistics South Africa.

Statistics South Africa. (2011). Mid-year population estimates (P0302). Pretoria: Statistics South Africa.

Stifel D., Rakotomanana, F. H., & Celada, E. (2007, June). *Assessing labor market conditions in Madagascar, 2001-2005* (African Region Working Paper Series No. 105). Washington, DC: World Bank.

Stock, J., & Yogo, M. (2004). Testing for weak instruments in linear IV regression? In *Identification and Inference in Econometric Models: Essays in Honor of Thomas J. Rothenberg*, ed. Donald Andrews, and James Stock. Cambridge: Cambridge University Press.

Stock, J., & Watson, M. (2003). *Introduction to econometrics*, Addison-Wesley.

Stock, J. H, Wright, J. H., & Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments, *Journal of Business and Economic Statistics,20*(4). 518 – 529.

Swedish International Development Cooperation Agency (SIDA). (2004). *The informal economy*. Stockholm, Sweden: SIDA.

Tannuri-Pianto, M., Pianto, D., & Arias, O. (2004). *Informal employment in Bolivia: A lost proposition?* (Econometric Society Latin American Meetings No. 149). New York, NY: Econometric Society.

Tansel, A. (2001). *Wage earners, self employed and gender in the informal sector in Turkey* (ERF Working Paper Series 0102). Cairo, Egypt: Economic Research Forum.

Tegoum, N. P. (2009). *Estimating the returns to education in Cameroon informal sector.* Presented at GLOBELICS 2009, 7th International Conference, 6-8 October, Dakar, Senegal.

Thomas, D. (1996). Education across generations in South Africa. *The American Economic Review, 86*(2), 330-334.

Trost, R. P., & Lee, L. (1984). Technical training and earnings: A polychotomous choice model with selectivity. *The Review of Economics and Statistics, 66*(1), 151-156.

Trostel, P., Walker, I., Woolley, P. (2002). Estimates of the economic return to schooling for 28 countries. *Labour Economics, 9*(1), 1-16.

Tueros, M. (1995). Education and informal labor markets. In M. Carnoy (Ed.), *Encyclopedia of*

economics of education (pp. 107-111). New York, NY: Pergamon.

United Nations Development Programme (UNDP). (2004). *Unleashing entrepreneurship: Making business work for the poor.* New York, NY: UNDP.

United Nations Educational, Scientific, and Cultural Organization (UNESCO). (August, 2010). World data on education: South Africa. Seventh edition, 2010/2011. Paris: UNESCO. Retrieved from http://www.ibe.unesco.org/fileadmin/user_upload/Publications/WDE/2010/pdf-versions/South_Africa.pdf

United States Internal Revenue Service. (2011). 1040 instructions. Washington, DC: United States Internatl Revenue Service.

Valodia, I., Lebani, L., Skinner, C., & Devey, R. (2006). Low-waged and informal employment in South Africa. *Transformation: Critical Perspectives on Southern Africa, 60*, 90-126.

van der Berg, S. (2002). *Education, poverty, and inequality in South Africa*. Paper to the conference of the Centre for the Study of African Economies on Economic Growth and Poverty in Africa, Oxford.

Vijverberg, W. P. (1993). Educational investments and returns for women and men in Côte d'lvoire. *Journal of Human Resources, 28*(40), 933-974.

Welsh.D, (2009). *The Rise and Fall of Apartheid.* Johannesburg and Cape Town: Jonathan Ball Publishers (PTY) LTD.

Wooldridge, J. M. (2002a). Unobserved heterogeneity and estimation of average partial effects (Michigan State University Working Paper). East Lansing, MI: Michigan State University.

Wooldridge, J.M. (2002b). Econometric analysis of cross section and panel data. London: the MIT Press.

World Bank. (1987). Education policies for Sub-Saharan Africa: Adjustment, revitalization and expansion (Report No. 6974). Washington, DC: World Bank.

World Bank. (2010). *Handbook on impact evaluation: Quantitative ,methods and practices.* Washington, DC: World Bank.

World Bank. (April 14, 2011). World development indicators and global development finance database. Retrieved from http://databank.worldbank.org/ddp/home.do.

World Bank. (May 4, 2012). World development indicators and global development finance database. Retrieved from http://databank.worldbank.org/ddp/home.do.

World Bank. (2010). Country Partnership Strategy Progress Report for the Republic of South Africa. Washington, DC: World Bank.

World Bank. (2012, April). South Africa country brief. Retrieved from http://web.worldbank.org/WBSITE/EXTERNAL/COUNTRIES/AFRICAEXT/SOUTHAFRICAEXTN/0,,menuPK:368086~pagePK:141132~piPK:141107~theSitePK:368057,00.html

APPENDICES

Table A.1 *Definitions of Variables Used in the Estimations*

Variable	Type of variable	Description
public	dummy	1 if working in the public sector, 0 otherwise
private, formal	dummy	1 if working in the formal private sector, 0 otherwise
private, informal	dummy	1 if working in the informal private sector, 0 otherwise
sector2	categorical	1 if unmployed/out of labor force, 2 if working in the public sector, 3 if working in the formal private sector, and 4 if working in the informal private sector (this variable is used especially for the choice model)
log hourly wage (lhwage)	continuous	Log of hourly wage. Hourly wage is calculated based on total salary of main occupation, salary period for the reported salary, and hours worked last week.
years of schooling	continuous	Years of schooling attained. It is constructed based on grade, certificates, diplomas, bachelors and higher degrees completed. (see Table A.2 for more details)
spouse years of schooling	continuous	Years of schooling attained by spouse
age	continuous	Age
experience	continuous	Experience in the current main occupation
female	dummy	1 if female, and 0 if male
married	dummy	1 if married, 0 otherwise
white	dummy	1 if White, 0 otherwise
black	dummy	1 if Black, 0 otherwise
colored	dummy	1 if Colored, 0 otherwise
asian	dummy	1 if Asian, 0 otherwise
permanent	dummy	1 if a worker has a permanent status at work, 0 otherwise
fixed	dummy	1 if a worker has a fixed period contract, 0 otherwise
temporary	dummy	1 if a worker has a temporary job, 0 otherwise
casual	dummy	1 if a worker has a casual job, 0 otherwise
seasonal	dummy	1 if a worker has a seasonal job, 0 otherwise
trade union	dummy	1 if a respondent is a member of trade union, 0 otherwise
inverse dependecy ratio	continuous	The number of working individuals devided by the number of household members.
training	dummy	1 if a respondent has been trained in skills that can be used for work, 0 otherwise
government job creation program	dummy	1 if a respondent work in any government job creation programme or project during the past six months

Table A.1 *(continued)*

Variable	Type of variable	Description
district	set of dummies	In South Africa there are 53 district councils (DCs) which consist of forty-seven (47) non-metropolitan areas and six (6) metropolitan areas.
industry	set of dummies	Types of industry included are (1) agriculture, hunting, forestry and fishing, (2) mining and quarrying, (3) manufacturing, (4) electricity, gas and water supply, (5) construction, (6) wholesale and retail trade (7) transport, storage and communication, (8) financial intermediation, insurance, real estate and business services, (9) community, social and personal services, (10) private households, and (11) 3xterior organisations and foreign government. "Unspecified" is categorized as a missing value.
occupation	set of dummies	Types of occupation included are (1) legislators, senior officials and managers, (2) professionals, (3) technical and associate professionals, (4) clerks, (5) service workers and shop and market sales workers, (6) skilled agricultural and fishery workers, (7) craft and related trades workers, (8) plant and machine operators and assemblers, (9) elementary occupation, and (10) domestic workers. "Unspecified" is categorized as a missing value.
language	set of dummies	Kinds of languages included are (1) Afrikaans, (2) English, (3) Isindebele / South Ndebele / North Ndebele, (4) Isixhosa / Xhosa, (5) Isizulu / Zulu, (6) Sepedi / Northern Sotho, (7) Sesotho / Southern Sotho / Sotho, (8) Setswana / Tswana, (9) Siswati / Swazi, (10) Tshivenda / Venda, (11) Xitsonga / Tsonga, and (12) Other. "Unspecified" is categorized as a missing value.
spouse public sector	dummy	1 if spouse is working in the public sector, 0 otherwise.
spouse formal private sector	dummy	1 if spouse is working in the formal private sector, 0 other wise.
spouse informal private sector	dummy	1 if spouse is working in the informal private sector, 0 otherwise.
househoud public sector	dummy	1 if at least one of the household members are working in the public sector, 0 otherwise.
household formal private sector	dummy	1 if at least one of the household members are working in the formal private sector sector, 0 otherwise.
household informal private sector	dummy	1 if at least one of the household members are working in the informal private sector, 0 otherwise.

Table A.2 *Years of Schooling and Corresponding Grade, Certificates, Diploma or Degree*

	years of schooling
No schooling	0
Grade R	0.5
Grade 1	1
Grade 2	2
Grade 3	3
Grade 4	4
Grade 5	5
Grade 6	6
Grade 7	7
Grade 8/NTCI	8
Grade 9/NTCII	9
Grade 10/NTCIII	10
Grade 11	11
Grade 12	12
Certificate with Grade 12	13
Diploma with Grade 12(not associate degree)	14
Bachelors Degree	15
Bacherlors Degree with Diploma	16
Honors Degree	17
Master's Degree and Doctorate	19

Notes: NTC means National Technical Certificate.

Table A.3 *Multinomial Employment Sector Choice Model Used for Selection Corrections, Marginal Effects*

	Public	Formal private	Informal private	Unemployed /out of labor force
	(1)	(2)	(3)	(4)
age	0.007***	0.084***	0.002***	-0.093***
	(0.001)	(0.002)	0.000	(0.002)
age squared/100	-0.007***	-0.100***	-0.001**	0.109***
	(0.001)	(0.003)	(0.001)	(0.003)
female*	-0.003**	-0.285***	0.026***	0.262***
	(0.001)	(0.006)	(0.003)	(0.006)
primary*	0.001	0.064***	-0.004	-0.060***
	(0.003)	(0.013)	(0.002)	(0.011)
lower secondary*	0.008*	0.103***	-0.013***	-0.099***
	(0.003)	(0.013)	(0.002)	(0.011)
upper secondary*	0.040***	0.216***	-0.034***	-0.222***
	(0.004)	(0.013)	(0.004)	(0.012)
certificates and	0.320***	0.005	-0.042***	-0.283***
associate degree*	(0.025)	(0.026)	(0.004)	(0.011)
bachelors and above*	0.341***	-0.004	-0.044***	-0.292***
	(0.028)	(0.029)	(0.004)	(0.011)
Black	0.015***	-0.380***	0.036***	0.329***
	(0.002)	(0.011)	(0.005)	(0.011)
Colored	0.038***	-0.232***	0.003	0.191***
	(0.004)	(0.013)	(0.005)	(0.013)
Asian	-0.002	-0.236***	-0.01	0.248***
	(0.003)	(0.019)	(0.006)	(0.020)
inverse dependency ratio	0.040***	1.112***	0.092***	-1.244***
	(0.004)	(0.025)	(0.009)	(0.027)
spouse formal private sector*	-0.003	0.004	0.004	-0.005
	(0.002)	(0.010)	(0.003)	(0.009)
spouse informal private sector*	-0.011***	-0.111***	0.010***	0.111***
	(0.002)	(0.015)	(0.003)	(0.015)
household formal private sector*	-0.016***	0.003	-0.030***	0.042***
	(0.001)	(0.008)	(0.003)	(0.007)
household informal private sector*	0.005***	-0.011	0.041***	-0.036***
	(0.001)	(0.007)	(0.004)	(0.006)
Pseudo R^2		0.3311		
N		65235		

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parenthesis.

Table A.4 *Details of the 2SLS Estimation with DMF Variant 1 Selection Correction, All Races*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.181***	0.251***	0.168***	0.145***	0.242***	0.143**	0.271**	0.253***	0.187***
	(0.040)	(0.020)	(0.031)	(0.043)	(0.024)	(0.049)	(0.089)	(0.039)	(0.043)
experience	0.002	0.034***	0.040***	0.009	0.046***	0.04	-0.003	0.012	0.035*
	(0.008)	(0.006)	(0.012)	(0.011)	(0.008)	(0.022)	(0.013)	(0.010)	(0.015)
experience^2	0.000*	0.000	0.000	0.000	0.000	0.000	0.001*	0.000*	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
female	-0.253***	-0.323***	-0.250***						
	(0.052)	(0.035)	(0.054)						
permanent	0.542***	0.202***	0.068	0.667***	0.198***	0.03	0.408***	0.204***	0.095
	(0.080)	(0.031)	(0.044)	(0.115)	(0.039)	(0.075)	(0.115)	(0.053)	(0.056)
trade union	0.268***	0.213***	0.242	0.221**	0.271***	0.085	0.224*	0.077	0.31
	(0.055)	(0.031)	(0.125)	(0.072)	(0.038)	(0.216)	(0.101)	(0.051)	(0.169)
Black	-0.208*	-0.633***	-1.295***	-0.402**	-0.578***	-1.762***	0.021	-0.709***	-0.892*
	(0.093)	(0.064)	(0.342)	(0.126)	(0.086)	(0.521)	(0.147)	(0.100)	(0.428)
Colored	-0.076	-0.367***	-1.036**	-0.238	-0.377***	-1.532**	0.234	-0.294*	-0.585
	(0.118)	(0.077)	(0.338)	(0.150)	(0.101)	(0.513)	(0.222)	(0.136)	(0.435)
Asian	-0.088	-0.299***	-0.043	-0.134	-0.250**	-0.637	-0.067	-0.331***	1.007*
	(0.115)	(0.065)	(0.472)	(0.121)	(0.090)	(0.619)	(0.273)	(0.095)	(0.431)
trnsp0	0.389**	0.330**	-0.047	0.348	0.371*	-0.096	0.564*	0.490*	-0.113
	(0.139)	(0.112)	(0.198)	(0.181)	(0.148)	(0.384)	(0.240)	(0.212)	(0.230)
trnsp1		1.399***	2.038***		1.493***	1.301		1.224*	2.582**
		(0.281)	(0.599)		(0.344)	(0.880)		(0.529)	(0.892)
trnsp2	0.189		-0.274	0.276		-0.336	0.271		-0.401
	(0.182)		(0.268)	(0.249)		(0.608)	(0.325)		(0.303)
trnsp3	-0.239	-0.354*		-0.378	-0.173		-0.225	-0.331	
	(0.307)	(0.171)		(0.380)	(0.211)		(0.550)	(0.282)	
lnp1	-0.086			-0.04			-0.245		
	(0.081)			(0.094)			(0.179)		
lnp2		-0.226***			-0.284**			-0.250**	
		(0.059)			(0.088)			(0.096)	
lnp3			0.001			0.031			0.052
			(0.077)			(0.126)			(0.096)
constant	0.445	-0.111	1.269**	1.004	-0.161	1.839*	-1.16	-0.156	0.529
	(0.744)	(0.269)	(0.489)	(0.849)	(0.330)	(0.844)	(1.555)	(0.485)	(0.620)
R squared	0.415	0.277	0.058	0.386	0.266	0.174	0.437	0.368	.
N	1312	3963	1140	792	2805	454	520	1158	686

Note: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in the parenthesis.
Selection terms are explained in Appendix 2.

Table A.5 *Details of the 2SLS Estimation with DMF Variant 1 Selection Correction, White*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.313*	0.492***	-0.364	0.135	0.389*	-0.169	0.551	0.785*	0.372
	(0.149)	(0.142)	(.)	(0.166)	(0.166)	(.)	(0.300)	(0.340)	(.)
experience	0.052	0.057*	-0.030	0.054	0.078	0.163	0.063	0.064	0.071
	(0.028)	(0.029)	(.)	(0.065)	(0.050)	(.)	(0.038)	(0.047)	(.)
experience^2	-0.001	-0.001	-0.0004	-0.001	-0.001	-0.007	-0.001	0.000	0.000
	(0.001)	(0.001)	(.)	(0.001)	(0.001)	(.)	(0.001)	(0.001)	(.)
female	-0.411**	-0.513**	-0.300						
	(0.145)	(0.176)	(.)						
permanent	0.657**	-0.126	0.739	1.422***	-0.317	0.000	0.410*	0.1	0.000
	(0.208)	(0.274)	(.)	(0.248)	(0.416)	(.)	(0.191)	(0.347)	(.)
trade union	0.174	0.293**	1.832	-0.055	0.256	3.061	0.163	0.314*	0.000
	(0.141)	(0.097)	(.)	(0.233)	(0.141)	(.)	(0.195)	(0.150)	(.)
trnsp0	0.756	1.531*	-0.035	1.766	1.53	0.000	1.071	2.315	0.000
	(1.089)	(0.681)	(.)	(2.083)	(0.916)	(.)	(1.406)	(1.424)	(.)
trnsp1		4.703*	-0.609		3.765	0.000		7.996	0.000
		(1.823)	(.)		(2.605)	(.)		(4.122)	(.)
trnsp2	1.196		4.843	3.023		0.000	0.737		0.000
	(1.469)		(.)	(2.092)		(.)	(2.940)		(.)
trnsp3	-1.534	-2.072		-0.197	-2.14		-3.768	-2.195	
	(1.552)	(1.384)		(2.030)	(2.005)		(3.041)	(1.822)	
lnp1	-0.293			-0.173			-0.507		
	(0.242)			(0.271)			(0.505)		
lnp2		-2.193*			-2.019			-3.549	
		(0.926)			(1.510)			(1.921)	
lnp3			-1.198			-0.308			-0.451
			(.)			(.)			(.)
constant	-1.413	-2.509	8.008	2.42	-1.396	4.872	-6.013	-6.63	-3.751
	(2.796)	(1.725)	(.)	(4.231)	(2.143)	(.)	(4.667)	(4.524)	(.)
R squared	0.291	0.003	1	0.43	0.139	1	.	.	1
N	127	383	11	63	212	6	64	171	5

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in the parenthesis. Selection terms are explained in Appendix 2.

Table A.6 *Details of the 2SLS Estimation with DMF Variant 1 Selection Correction, Black*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.196***	0.230***	0.157***	0.121*	0.230***	0.119*	0.404*	0.222***	0.188***
	(0.058)	(0.025)	(0.035)	(0.055)	(0.029)	(0.053)	(0.185)	(0.050)	(0.050)
experience	-0.001	0.043***	0.047***	-0.001	0.060***	0.05	0.023	0.016	0.037*
	(0.009)	(0.009)	(0.014)	(0.012)	(0.011)	(0.026)	(0.027)	(0.016)	(0.017)
experience^2	0.001**	0.000	-0.000*	0.000	-0.000**	-0.001	0.001	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
female	-0.197**	-0.352***	-0.244***						
	(0.071)	(0.046)	(0.065)						
permanent	0.522***	0.189***	0.028	0.713***	0.175***	-0.063	0.332*	0.222**	0.093
	(0.097)	(0.037)	(0.050)	(0.133)	(0.045)	(0.087)	(0.165)	(0.071)	(0.063)
trade union	0.226**	0.264***	0.315*	0.200*	0.327***	0.217	0.070	0.073	0.369
	(0.073)	(0.040)	(0.146)	(0.092)	(0.048)	(0.238)	(0.182)	(0.073)	(0.210)
trnsp0	0.267	0.550***	0.170	0.224	0.582**	0.373	0.664	0.614	-0.200
	(0.180)	(0.146)	(0.259)	(0.222)	(0.184)	(0.405)	(0.372)	(0.342)	(0.307)
trnsp1		1.566***	2.252**		1.876***	1.266		0.872	3.064**
		(0.390)	(0.745)		(0.475)	(1.098)		(0.809)	(1.093)
trnsp2	-0.129		-0.147	-0.158		0.056	0.375		-0.549
	(0.222)		(0.331)	(0.279)		(0.612)	(0.513)		(0.375)
trnsp3	-0.347	0.102		-0.285	0.235		-0.617	0.232	
	(0.361)	(0.177)		(0.405)	(0.223)		(0.825)	(0.328)	
lnp1	-0.092			0.029			-0.518		
	(0.130)			(0.135)			(0.398)		
lnp2		-0.240***			-0.296**			-0.237	
		(0.072)			(0.107)			(0.122)	
lnp3			-0.069			-0.138			0.122
			(0.106)			(0.149)			(0.138)
constant	-0.197	-0.369	0.083	0.807	-0.494	0.355	-3.497	-0.366	-0.363
	(0.998)	(0.278)	(0.372)	(0.992)	(0.341)	(0.727)	(3.187)	(0.506)	(0.490)
R squared	0.367	0.125	-	0.364	0.105	0.064	0.248	0.209	-
N	865	2377	853	524	1782	345	341	595	508

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in the parenthesis. Selection terms are explained in Appendix 2.

Table A.7 *Details of the 2SLS Estimation with DMF Variant 1 Selection Correction, Colored*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.135	0.342***	0.210**	0.227*	0.338***	0.229	-0.047	0.338**	0.177*
	(0.071)	(0.059)	(0.066)	(0.105)	(0.070)	(0.121)	(0.120)	(0.106)	(0.086)
experience	0.022	0.059***	0.024	0.078*	0.071***	0.007	0.005	0.037	0.024
	(0.021)	(0.015)	(0.026)	(0.036)	(0.019)	(0.045)	(0.032)	(0.022)	(0.036)
experience^2	0.000	0.000	0.000	-0.001	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)
female	-0.238*	-0.309***	-0.193	0.236	0.329**	0.156	0.527*	0.164	0.044
	(0.109)	(0.084)	(0.107)	(0.277)	(0.108)	(0.177)	(0.248)	(0.107)	(0.123)
permanent	0.488**	0.240***	0.117	0.600***	0.194*	-0.12	0.594**	-0.058	-0.126
	(0.179)	(0.072)	(0.095)	(0.177)	(0.087)	(0.629)	(0.202)	(0.124)	(0.264)
trade union	0.466***	0.117	-0.100						
	(0.120)	(0.071)	(0.264)						
trnsp0	0.570	-0.106	-1.487**	0.611	-0.282	-2.680*	0.204	0.506	-1.032*
	(0.361)	(0.303)	(0.505)	(0.485)	(0.434)	(1.174)	(0.779)	(0.503)	(0.508)
trnsp1		2.293**	1.642		2.156*	1.911		2.812	0.822
		(0.795)	(1.249)		(0.959)	(1.955)		(1.547)	(1.725)
trnsp2	0.480		-2.451**	1.376		-4.771*	-0.853		-1.800*
	(0.495)		(0.841)	(0.820)		(1.912)	(0.880)		(0.760)
trnsp3	0.488	-1.121*		-0.656	-0.923		1.742*	-0.978	
	(0.720)	(0.555)		(1.376)	(0.711)		(0.721)	(0.770)	
lnp1	-0.035			-0.366			0.507		
	(0.195)			(0.296)			(0.428)		
lnp2		-0.313			-0.41			-0.373	
		(0.207)			(0.326)			(0.302)	
lnp3			0.363*			0.680*			0.161
			(0.173)			(0.324)			(0.172)
constant	1.057	-1.952**	-1.441	-0.73	-2.248*	-2.813	3.389	-1.507	-1.195
	(1.140)	(0.705)	(0.970)	(1.869)	(0.912)	(2.031)	(1.749)	(1.103)	(1.076)
R squared	0.515	.	.	0.294	.	0.07	0.663	.	.
N	276	1035	272	174	700	100	102	335	172

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in the parenthesis. Selection terms are explained in Appendix 2.

Table A.8 *Details of the 2SLS Estimation with DMF Variant 1 Selection Correction, Asian*

Total	Total			Male			Female		
	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal	Public	Private, formal	Private, informal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
years of schooling	0.179	0.412		0.116	0.292		1.443	0.634	
	(0.129)	(0.327)		(0.148)	(0.306)		(1.239)	(0.677)	
experience	-0.052	0.027		-0.067	0.014		-0.449	0.014	
	(0.043)	(0.041)		(0.070)	(0.042)		(0.408)	(0.091)	
experience^2	0.001	0		0.001	0		0.012	0.001	
	(0.001)	(0.001)		(0.001)	(0.001)		(0.010)	(0.001)	
female	-0.541	-0.583*							
	(0.421)	(0.284)							
permanent	0.722***	0.218		0.650***	0.28		0	0.125	
	(0.177)	(0.156)		(0.163)	(0.199)		(.)	(0.261)	
trade union	-0.088	0.038		-0.038	-0.078		-0.228	0.499	
	(0.310)	(0.112)		(0.408)	(0.155)		(0.791)	(0.303)	
trnsp0	-0.106	0.871		0.482	1.295		-13.873	0.436	
	(1.109)	(0.659)		(0.992)	(0.936)		(13.768)	(1.767)	
trnsp1		1.451			0.7			2.494	
		(2.741)			(2.671)			(6.114)	
trnsp2	-0.847			-0.266			-28.241		
	(1.437)			(1.533)			(24.875)		
trnsp3	0.052	-2.731		-0.209	-1.729		-12.932	-5.058	
	(2.757)	(3.191)		(3.233)	(4.637)		(15.806)	(4.279)	
lnp1	0.166			0.208			2.183		
	(0.189)			(0.309)			(1.937)		
lnp2		-0.442			-0.672			-0.159	
		(0.460)			(0.557)			(1.709)	
lnp3									
constant	0.936	-2.325		2.628	-0.72		-35.221	-5.903	
	(2.809)	(4.120)		(3.259)	(3.937)		(34.587)	(8.722)	
R squared	0.413	0.074		0.574	0.203		0.635	.	
N	44	168	4	31	111	3	13	57	

Notes: * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors are in the parenthesis. Selection terms are explained in Appendix 2.

Table A.9 *Details of the Probit and IV Probit Estimation, Public vs. Formal Private, All Races*

	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
yschool	0.013***	0.010***	0.041***	0.014***	0.009***	0.054***	0.014***	0.009***	0.038***
	(0.001)	(0.001)	(0.004)	(0.001)	(0.001)	(0.005)	(0.001)	(0.001)	(0.004)
female*	0.035***			0.049***			0.040***		
	(0.004)			(0.007)			(0.007)		
age	0.009***	0.010***	0.016***	0.013***	0.012***	0.025***	0.016***	0.015***	0.021***
	(0.001)	(0.002)	(0.004)	(0.002)	(0.002)	(0.007)	(0.002)	(0.002)	(0.006)
age^2/100	-0.007***	-0.009***	-0.009*	-0.012***	-0.011***	-0.018*	-0.014***	-0.014***	-0.017*
	(0.001)	(0.002)	(0.004)	(0.002)	(0.003)	(0.009)	(0.002)	(0.003)	(0.008)
Black*	0.077***	0.083***	0.107**	0.095***	0.100***	0.101	0.140***	0.121***	0.220***
	(0.017)	(0.023)	(0.040)	(0.025)	(0.028)	(0.068)	(0.010)	(0.011)	(0.023)
Colored*	0.090***	0.105***	0.139***	0.108***	0.106***	0.164***	0.112***	0.117***	0.132***
	(0.010)	(0.015)	(0.026)	(0.014)	(0.017)	(0.034)	(0.012)	(0.014)	(0.030)
Asian*	0.038**	0.049**	0.063	0.031*	0.050**	-0.002	0.003	0.019	-0.041
	(0.012)	(0.016)	(0.038)	(0.015)	(0.018)	(0.052)	(0.012)	(0.014)	(0.035)
married*	0.002	-0.007	0.061***						
	(0.003)	(0.004)	(0.016)						
inverse dependecy	-0.030***	-0.037***	-0.058***	-0.037***	-0.038***	-0.058*	-0.038***	-0.039***	-0.057*
ratio	(0.005)	(0.007)	(0.015)	(0.007)	(0.008)	(0.026)	(0.008)	(0.009)	(0.024)
training*	0.002	0.003	0.005	0.000	0.001	-0.003	0.003	0.005	0.001
	(0.003)	(0.005)	(0.013)	(0.005)	(0.006)	(0.024)	(0.007)	(0.007)	(0.022)
government job	0.113***	0.095**	0.271***	0.150***	0.135*	0.315***	0.179***	0.142*	0.328***
creation program*	(0.023)	(0.033)	(0.048)	(0.041)	(0.053)	(0.078)	(0.049)	(0.061)	(0.087)
spouse	-0.009	0.001	-0.061***	0.010	0.026	-0.042	0.028	0.041	-0.027
public sector*	(0.006)	(0.009)	(0.018)	(0.015)	(0.019)	(0.052)	(0.021)	(0.025)	(0.051)
spouse formal	0.004	0.027**	-0.058***	-0.007	0.018	-0.113***	-0.010	0.013	-0.103***
sector*	(0.005)	(0.010)	(0.015)	(0.008)	(0.010)	(0.029)	(0.009)	(0.011)	(0.024)
household	0.104***	0.122***	0.170***	0.086***	0.083**	0.143*	0.080**	0.072*	0.135
public sector*	(0.013)	(0.019)	(0.026)	(0.024)	(0.028)	(0.064)	(0.027)	(0.029)	(0.071)
household	-0.032***	-0.034***	-0.083***	-0.031***	-0.026***	-0.075*	-0.028***	-0.026***	-0.030
formal sector*	(0.004)	(0.006)	(0.014)	(0.006)	(0.006)	(0.030)	(0.007)	(0.006)	(0.031)
Pseudo R^2	0.185	0.148	0.221	0.170	0.135	0.233			
N	17786	10645	7106	9774	6478	3296	8474	5636	2838

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parenthesis. Language dummies are also included in the ordinary probit estimations. * after the name of the variable means that dy/dx is for discrete change of dummy variable from 0 to 1.

Table A.10 *Details of the Probit and IV Probit Estimation, Public vs. Formal Private, Whites*

	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
yschool	0.013***	0.003	0.035***	0.011**	0.002	0.037***	0.007	-0.010	0.036*
	(0.004)	(0.005)	(0.008)	(0.004)	(0.005)	(0.009)	(0.007)	(0.008)	(0.016)
female*	0.080***			0.085***			0.081***		
	(0.019)			(0.022)			(0.022)		
age	0.037***	0.049***	0.030**	0.029***	0.041***	0.017	0.030***	0.042***	0.017
	(0.007)	(0.011)	(0.010)	(0.007)	(0.011)	(0.013)	(0.007)	(0.010)	(0.013)
age^2/100	-0.039***	-0.055***	-0.028*	-0.031***	-0.046***	-0.013	-0.032***	-0.048***	-0.013
	(0.008)	(0.012)	(0.012)	(0.008)	(0.012)	(0.015)	(0.008)	(0.011)	(0.016)
married*	-0.043	-0.056	0.000						
	(0.025)	(0.035)	(0.053)						
inverse dependecy	0.011	0.019	-0.014	0.007	0.005	-0.003	0.011	0.006	0.005
ratio	(0.023)	(0.033)	(0.040)	(0.024)	(0.034)	(0.047)	(0.025)	(0.034)	(0.050)
training*	-0.012	-0.011	-0.026	-0.009	-0.017	-0.006	-0.010	-0.014	-0.020
	(0.020)	(0.027)	(0.036)	(0.020)	(0.025)	(0.045)	(0.020)	(0.025)	(0.046)
government job	0.136	0.238	0.057	0.050	0.151	-0.079	0.087	0.230	-0.076
creation program*	(0.132)	(0.192)	(0.202)	(0.117)	(0.185)	(0.181)	(0.135)	(0.221)	(0.192)
spouse	-0.046	-0.076	0.022	0.028	-0.011	0.218	0.056	0.022	0.194
public sector*	(0.038)	(0.042)	(0.103)	(0.075)	(0.072)	(0.247)	(0.088)	(0.088)	(0.248)
spouse formal	0.024	0.030	-0.016	-0.041	-0.027	-0.101	-0.054*	-0.026	-0.154*
sector*	(0.035)	(0.047)	(0.061)	(0.029)	(0.036)	(0.071)	(0.028)	(0.036)	(0.068)
household	0.204***	0.301***	0.072	0.084	0.166	-0.071	0.051	0.121	-0.065
public sector*	(0.056)	(0.075)	(0.090)	(0.085)	(0.115)	(0.155)	(0.082)	(0.112)	(0.164)
household	-0.072**	-0.067*	-0.095*	-0.003	-0.004	0.011	0.013	-0.004	0.093
formal sector*	(0.023)	(0.031)	(0.042)	(0.034)	(0.039)	(0.096)	(0.038)	(0.039)	(0.118)
Pseudo R^2	0.077	0.079	0.091	0.057	0.053	0.079			
N	2386	1349	1036	1813	1051	760	1773	1033	740

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parenthesis. Language dummies are also included in the ordinary probit estimations. * after the name of the variable means that dy/dx is for discrete change of dummy variable from 0 to 1.

Table A.11 *Details of the Probit and IV Probit Estimation, Public vs. Formal Private, Blacks*

	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
yschool	0.018***	0.017**	0.034**	0.018**	0.018***	0.021	0.028***	0.021***	0.057***
	(0.005)	(0.005)	(0.013)	(0.005)	(0.005)	(0.016)	(0.003)	(0.003)	(0.008)
female*	0.055***			0.076***			0.096***		
	(0.014)			(0.022)			(0.017)		
age	0.010**	0.011**	0.013*	0.012**	0.016*	0.009	0.025***	0.024***	0.033*
	(0.003)	(0.004)	(0.006)	(0.004)	(0.006)	(0.007)	(0.005)	(0.005)	(0.013)
age^2/100	-0.006*	-0.008*	-0.006	-0.009*	-0.012*	-0.005	-0.020***	-0.020**	-0.025
	(0.002)	(0.003)	(0.004)	(0.004)	(0.006)	(0.006)	(0.006)	(0.006)	(0.016)
married*	0.007	-0.007	0.053*						
	(0.005)	(0.008)	(0.022)						
inverse dependecy	-0.045***	-0.058**	-0.059*	-0.053**	-0.078**	-0.034	-0.110***	-0.110***	-0.147**
ratio	(0.013)	(0.020)	(0.026)	(0.018)	(0.027)	(0.028)	(0.021)	(0.024)	(0.053)
training*	-0.001	0.003	-0.008	-0.001	0.003	-0.005	0.012	0.015	0.01
	(0.005)	(0.008)	(0.012)	(0.008)	(0.013)	(0.012)	(0.017)	(0.019)	(0.047)
government job	0.129***	0.118*	0.232***	0.199**	0.244**	0.229	0.315***	0.253*	0.445***
creation program*	(0.037)	(0.051)	(0.070)	(0.064)	(0.092)	(0.130)	(0.070)	(0.100)	(0.074)
spouse	-0.016	0.005	-0.056*	-0.003	0.049	-0.033	0.040	0.132	-0.156
public sector*	(0.010)	(0.018)	(0.027)	(0.021)	(0.050)	(0.029)	(0.054)	(0.079)	(0.092)
spouse formal	0.000	0.032	-0.044*	-0.016	0.026	-0.037	-0.029	0.025	-0.189***
sector*	(0.009)	(0.020)	(0.022)	(0.012)	(0.026)	(0.032)	(0.025)	(0.034)	(0.050)
household	0.149***	0.160***	0.194***	0.132**	0.136*	0.146	0.175**	0.095	0.376***
public sector*	(0.033)	(0.043)	(0.055)	(0.050)	(0.066)	(0.096)	(0.065)	(0.070)	(0.111)
household	-0.044**	-0.052**	-0.065*	-0.038*	-0.057*	-0.022	-0.056**	-0.066***	0.008
formal sector*	(0.014)	(0.019)	(0.030)	(0.015)	(0.023)	(0.021)	(0.020)	(0.019)	(0.072)
Pseudo R^2	0.223	0.174	0.254	0.210	0.160	0.276			
N	11324	6901	4397	5601	3893	1708	4437	3116	1321

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parenthesis. Language dummies are also included in the ordinary probit estimations. * after the name of the variable means that dy/dx is for discrete change of dummy variable from 0 to 1.

Table A.12 *Details of the Probit and IV Probit Estimation, Public vs. Formal Private, Coloreds*

	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
yschool	0.025***	0.020***	0.032***	0.025***	0.013***	0.052***	0.032***	0.023***	0.050***
	(0.004)	(0.004)	(0.006)	(0.004)	(0.004)	(0.008)	(0.006)	(0.006)	(0.013)
female*	0.004			0.010			0.006		
	(0.011)			(0.017)			(0.022)		
age	0.008*	0.014**	-0.003	0.019**	0.013*	0.021	0.025**	0.022**	0.031
	(0.004)	(0.005)	(0.006)	(0.007)	(0.006)	(0.020)	(0.008)	(0.009)	(0.021)
age^2/100	0.008*	0.014**	-0.003	-0.016	-0.012	-0.011	-0.020*	-0.018	-0.023
	(0.004)	(0.005)	(0.006)	(0.008)	(0.007)	(0.024)	(0.010)	(0.010)	(0.026)
married*	-0.008	-0.023	0.050						
	(0.015)	(0.018)	(0.034)						
inverse dependecy	-0.052**	-0.092**	0.011	-0.065*	-0.075**	0.016	-0.079*	-0.115**	(0.018)
ratio	(0.020)	(0.029)	(0.034)	(0.027)	(0.028)	(0.070)	(0.036)	(0.040)	(0.079)
training*	0.061**	0.026	0.102**	0.037	0.014	0.083	0.027	0.016	0.053
	(0.022)	(0.024)	(0.038)	(0.026)	(0.021)	(0.064)	(0.033)	(0.036)	(0.069)
government job	0.356**	0.181	0.524***	0.097	-0.004	0.334	0.139	0.009	0.334
creation program*	(0.113)	(0.140)	(0.142)	(0.165)	(0.097)	(0.279)	(0.195)	(0.192)	(0.277)
spouse	0.033	0.036	0.001	0.061	0.042	0.141	0.066	0.051	0.162
public sector*	(0.032)	(0.041)	(0.047)	(0.056)	(0.048)	(0.148)	(0.068)	(0.072)	(0.165)
spouse formal	0.026	0.062	-0.044	0.021	0.030	-0.043	0.025	0.048	-0.044
sector*	(0.023)	(0.034)	(0.029)	(0.029)	(0.026)	(0.077)	(0.038)	(0.042)	(0.091)
household	0.086**	0.126**	0.044	0.060	0.083	-0.075	0.082	0.147	-0.103
public sector*	(0.027)	(0.041)	(0.037)	(0.051)	(0.055)	(0.109)	(0.063)	(0.078)	(0.117)
household	-0.073***	-0.065***	-0.087***	-0.074***	-0.034*	-0.177**	-0.104***	-0.073**	-0.201***
formal sector*	(0.015)	(0.019)	(0.025)	(0.020)	(0.016)	(0.059)	(0.022)	(0.024)	(0.058)
Pseudo R^2	0.164	0.165	0.183	0.156	0.149	0.204			
N	3440	1976	1456	1905	1219	685	1835	1194	641

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parenthesis. Language dummies are also included in the ordinary probit estimations. * after the name of the variable means that dy/dx is for discrete change of dummy variable from 0 to 1.

Table A.13 *Details of the Probit and IV Probit Estimation, Public vs. Formal Private, Asians*

	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
yschool	0.039***	0.036**	0.037*	0.044***	0.035**	0.121**	0.017	0.023***	0.085
	(0.010)	(0.013)	(0.017)	(0.011)	(0.011)	(0.039)	(0.017)	(0.006)	(0.060)
female*	0.000			-0.014			-0.018		
	(0.037)			(0.043)			(0.038)		
age	0.024	0.019	0.024	0.045*	0.064**	0.032	0.032*	0.022**	0.016
	(0.014)	(0.019)	(0.018)	(0.018)	(0.023)	(0.062)	(0.015)	(0.009)	(0.067)
age^2/100	-0.030	-0.023	-0.033	-0.052*	-0.072**	-0.049	-0.037*	-0.018	-0.030
	(0.017)	(0.022)	(0.024)	(0.022)	(0.027)	(0.079)	(0.018)	(0.010)	(0.088)
married*	-0.014	-0.056	0.291*						
	(0.052)	(0.072)	(0.141)						
inverse dependecy	-0.009	0.029	-0.047	0.053	0.086	-0.105	0.045	-0.115**	-0.064
ratio	(0.057)	(0.083)	(0.063)	(0.071)	(0.086)	(0.207)	(0.067)	(0.040)	(0.204)
training*	-0.025	-0.034	-0.025	-0.028	-0.020	-0.128	-0.027	0.016	-0.134
	(0.046)	(0.063)	(0.045)	(0.052)	(0.059)	(0.164)	(0.046)	(0.036)	(0.163)
government job	0.347	0.392		0.322	0.318		0.448	0.009	
creation program*	(0.317)	(0.303)		(0.322)	(0.320)		(0.321)	(0.192)	
spouse	0.030	0.182	-0.079	0.020	0.228	-0.329*	0.085	0.051	-0.316
public sector*	(0.104)	(0.188)	(0.043)	(0.121)	(0.227)	(0.151)	(0.143)	(0.072)	(0.169)
spouse formal	-0.038	0.141	-0.084	-0.023	0.155	-0.342*	-0.006	0.048	-0.346*
sector*	(0.046)	(0.115)	(0.047)	(0.062)	(0.130)	(0.151)	(0.068)	(0.042)	(0.162)
household	0.102	0.095	0.160	0.118	0.041	0.494*	0.072	0.147	0.412
public sector*	(0.092)	(0.118)	(0.158)	(0.133)	(0.136)	(0.205)	(0.116)	(0.078)	(0.249)
household	-0.035	-0.115	0.113	-0.045	-0.105*	0.503**	-0.037	-0.073**	0.511**
formal sector*	(0.044)	(0.059)	(0.072)	(0.055)	(0.044)	(0.183)	(0.053)	(0.024)	(0.195)
Pseudo R^2	0.091	0.096	0.194	0.125	0.147	0.316			
N	612	396	209	441	298	137	429	1194	136

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parenthesis. Language dummies are also included in the ordinary probit estimations. * after the name of the variable means that dy/dx is for discrete change of dummy variable from 0 to 1.

Table A.14 *Details of the Probit and IV Probit Estimation, Public vs. Informal Private, All Races*

	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
yschool	0.074***	0.049***	0.094***	0.067***	0.049***	0.095***	0.071***	0.048***	0.092***
	(0.002)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.005)	(0.006)
female*	-0.099***			-0.176***			-0.174***		
	(0.011)			(0.016)			(0.017)		
age	0.034***	0.041***	0.020***	0.039***	0.055***	0.028**	0.040***	0.053***	0.025*
	(0.004)	(0.005)	(0.005)	(0.006)	(0.009)	(0.010)	(0.007)	(0.009)	(0.010)
age^2/100	-0.030***	-0.041***	-0.012	-0.036***	-0.055***	-0.022	-0.037***	-0.053***	-0.020
	(0.004)	(0.006)	(0.007)	(0.007)	(0.010)	(0.012)	(0.008)	(0.010)	(0.012)
Black*	-0.075	-0.003	-0.160*	-0.079	0.041	-0.179*	-0.131***	-0.061	-0.165***
	(0.044)	(0.053)	(0.067)	(0.059)	(0.076)	(0.087)	(0.034)	(0.043)	(0.048)
Colored*	0.006	0.060	-0.071	-0.021	0.034	-0.050	0.013	0.063	-0.011
	(0.030)	(0.037)	(0.046)	(0.037)	(0.050)	(0.057)	(0.038)	(0.049)	(0.055)
Asian*	-0.019	0.007	0.009	-0.111	-0.044	-0.015	-0.073	-0.048	0.045
	(0.056)	(0.064)	(0.099)	(0.071)	(0.088)	(0.134)	(0.062)	(0.069)	(0.126)
married*	0.025	0.076***	-0.037						
	(0.014)	(0.019)	(0.019)						
inverse dependecy	-0.088***	-0.066**	-0.096***	-0.091***	-0.090**	-0.105**	-0.208***	-0.207***	-0.184***
ratio	(0.017)	(0.021)	(0.026)	(0.025)	(0.034)	(0.040)	(0.031)	(0.041)	(0.045)
training*	0.079***	0.087***	0.052*	0.055*	0.079*	0.014	0.049	0.092**	-0.035
	(0.018)	(0.024)	(0.026)	(0.023)	(0.032)	(0.038)	(0.027)	(0.034)	(0.039)
government job	0.383***	0.157*	0.493***	0.276***	0.151	0.446***	0.356***	0.162	0.578***
creation program*	(0.031)	(0.066)	(0.042)	(0.042)	(0.091)	(0.054)	(0.050)	(0.107)	(0.050)
spouse	-0.024	-0.059	0.042	0.095	0.058	0.132	0.149*	0.134	0.158
public sector*	(0.032)	(0.037)	(0.046)	(0.049)	(0.077)	(0.075)	(0.061)	(0.084)	(0.098)
spouse formal	-0.004	0.078*	-0.038	0.014	0.121*	-0.079	0.030	0.110	-0.036
sector*	(0.025)	(0.040)	(0.035)	(0.036)	(0.052)	(0.051)	(0.040)	(0.057)	(0.055)
household	0.309***	0.350***	0.281***	0.200***	0.244***	0.202**	0.190***	0.187*	0.211*
public sector*	(0.021)	(0.033)	(0.029)	(0.039)	(0.063)	(0.066)	(0.055)	(0.076)	(0.092)
household	0.035	0.070**	0.012	0.050	0.067	0.067	0.077*	0.101*	0.076
formal sector*	(0.018)	(0.026)	(0.025)	(0.031)	(0.044)	(0.050)	(0.037)	(0.047)	(0.058)
Pseudo R^2	0.312	0.263	0.378	0.304	0.223	0.412			
N	11272	4999	6233	5829	2876	2953	4852	2546	2306

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parenthesis. Language dummies are also included in the ordinary probit estimations. * after the name of the variable means that dy/dx is for discrete change of dummy variable from 0 to 1.

Table A.15 *Details of the Probit and IV Probit Estimation, Public vs. Informal Private, Whites*

	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
yschool	0.057***	0.074***	0.026*	0.049***	0.060**	0.045*	0.047*	0.044	0.058*
	(0.012)	(0.017)	(0.012)	(0.013)	(0.019)	(0.020)	(0.019)	(0.028)	(0.029)
female*	0.089			0.008			-0.001		
	(0.046)			(0.049)			(0.056)		
age	0.048**	0.063*	0.009	0.036	0.040	0.023	0.039	0.043	0.027
	(0.016)	(0.026)	(0.013)	(0.019)	(0.033)	(0.025)	(0.022)	(0.037)	(0.030)
age^2/100	-0.055**	-0.075*	-0.008	-0.039	-0.045	-0.021	-0.043	-0.048	-0.024
	(0.019)	(0.030)	(0.015)	(0.022)	(0.037)	(0.028)	(0.025)	(0.040)	(0.034)
married*	0.021	0.170	-0.123						
	(0.065)	(0.091)	(0.082)						
inverse dependecy	-0.011	0.027	-0.036	-0.033	-0.030	-0.021	-0.056	-0.065	-0.038
ratio	(0.066)	(0.104)	(0.051)	(0.073)	(0.119)	(0.095)	(0.082)	(0.127)	(0.114)
training*	-0.114	-0.065	-0.108	-0.071	0.017	-0.166	-0.077	0.022	-0.205
	(0.060)	(0.086)	(0.070)	(0.066)	(0.094)	(0.102)	(0.071)	(0.100)	(0.109)
government job creation program*									
spouse	-0.360*	-0.487***	0.078	0.032	-0.706***	0.200*	0.093	-0.659***	0.265**
public sector*	(0.174)	(0.098)	(0.048)	(0.178)	(0.088)	(0.084)	(0.188)	(0.062)	(0.091)
spouse formal	0.058	0.108	0.058	0.053	0.111	-0.022	0.079	0.154	-0.022
sector*	(0.081)	(0.133)	(0.036)	(0.086)	(0.114)	(0.186)	(0.096)	(0.118)	(0.213)
household	0.315***	0.513***	0.036	0.146	0.294***	-0.340	0.153	0.341***	-0.355
public sector*	(0.065)	(0.098)	(0.082)	(0.104)	(0.088)	(0.386)	(0.145)	(0.062)	(0.368)
household	0.018	0.065	-0.063	0.037	0.030	0.088	0.057	0.030	0.131
formal sector*	(0.071)	(0.105)	(0.077)	(0.084)	(0.116)	(0.134)	(0.094)	(0.127)	(0.157)
Pseudo R^2	0.065	0.121	0.111	0.107	0.141	0.116			
N	570	296	274	423	216	207	420	217	203

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parenthesis. Language dummies are also included in the ordinary probit estimations. * after the name of the variable means that dy/dx is for discrete change of dummy variable from 0 to 1.

Table A.16 *Details of the Probit and IV Probit Estimation, Public vs. Informal Private, Blacks*

	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
yschool	0.047***	0.032***	0.046*	0.051***	0.044***	0.019	0.064***	0.042***	0.062***
	(0.010)	(0.010)	(0.020)	(0.011)	(0.007)	(0.019)	(0.004)	(0.005)	(0.006)
female*	-0.061***			-0.112***			-0.154***		
	(0.016)			(0.032)			(0.016)		
age	0.023***	0.027**	0.013*	0.032***	0.052***	0.006	0.040***	0.051***	0.017*
	(0.006)	(0.009)	(0.006)	(0.009)	(0.012)	(0.006)	(0.007)	(0.009)	(0.008)
age^2/100	-0.020***	-0.026**	-0.009	-0.029**	-0.051***	-0.004	-0.037***	-0.052***	-0.012
	(0.005)	(0.009)	(0.005)	(0.009)	(0.013)	(0.005)	(0.008)	(0.010)	(0.009)
married*	0.018	0.051**	-0.012						
	(0.010)	(0.019)	(0.011)						
inverse dependecy	-0.041**	-0.028	-0.037	-0.037	-0.038	-0.015	-0.177***	-0.176***	-0.119***
ratio	(0.015)	(0.017)	(0.022)	(0.023)	(0.035)	(0.018)	(0.034)	(0.043)	(0.035)
training*	0.052**	0.059*	0.026	0.034	0.054	0.005	0.039	0.071*	-0.014
	(0.017)	(0.024)	(0.019)	(0.023)	(0.035)	(0.011)	(0.030)	(0.036)	(0.031)
government job	0.378***	0.107	0.551***	0.321***	0.131	0.320	0.406***	0.129	0.667***
creation program*	(0.055)	(0.061)	(0.081)	(0.069)	(0.096)	(0.182)	(0.064)	(0.107)	(0.068)
spouse	-0.029	-0.030	-0.006	0.035	0.048	0.006	0.127	0.180	0.052
public sector*	(0.020)	(0.027)	(0.023)	(0.050)	(0.092)	(0.020)	(0.077)	(0.112)	(0.080)
spouse formal	-0.016	0.036	-0.021	-0.010	0.102	-0.016	0.020	0.098	-0.023
sector*	(0.018)	(0.037)	(0.020)	(0.032)	(0.069)	(0.019)	(0.047)	(0.072)	(0.040)
household	0.269***	0.289***	0.211**	0.235***	0.272**	0.088	0.201**	0.152	0.201*
public sector*	(0.041)	(0.060)	(0.067)	(0.062)	(0.088)	(0.079)	(0.074)	(0.102)	(0.100)
household	0.016	0.039	0.007	0.035	0.037	0.018	0.067	0.077	0.054
formal sector*	(0.014)	(0.024)	(0.014)	(0.032)	(0.052)	(0.023)	(0.044)	(0.058)	(0.051)
Pseudo R^2	0.310	0.261	0.371	0.295	0.214	0.401			
N	9095	3911	5147	4415	2130	2285	3499	1818	1681

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parenthesis. Language dummies are also included in the ordinary probit estimations. * after the name of the variable means that dy/dx is for discrete change of dummy variable from 0 to 1.

Table A.17 *Details of the Probit and IV Probit Estimation, Public vs. Informal Private, Coloreds*

	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
yschool	0.065***	0.049***	0.082***	0.059***	0.045***	0.082***	0.071***	0.131***	0.077***
	(0.005)	(0.007)	(0.007)	(0.007)	(0.008)	(0.010)	(0.009)	(0.029)	(0.014)
female*	-0.126***			-0.216***			-0.192***		
	(0.031)			(0.040)			(0.039)		
age	0.009	0.034*	-0.023	0.024	0.051*	0.007	0.026	0.108	0.011
	(0.010)	(0.014)	(0.015)	(0.016)	(0.023)	(0.029)	(0.018)	.	(0.026)
age^2/100	-0.004	-0.035*	0.034	-0.021	-0.050	-0.004	-0.022	-0.099	-0.008
	(0.012)	(0.017)	(0.018)	(0.019)	(0.026)	(0.034)	(0.021)	(0.065)	(0.031)
married*	0.019	0.060	-0.058						
	(0.040)	(0.062)	(0.057)						
inverse dependecy	-0.256***	-0.325***	-0.155*	-0.309***	-0.429***	-0.226*	-0.336***	-0.998***	-0.253*
ratio	(0.053)	(0.075)	(0.078)	(0.074)	(0.101)	(0.108)	(0.076)	(0.270)	(0.102)
training*	0.178***	0.180*	0.192**	0.208***	0.293**	0.214*	0.206**	0.728*	0.164
	(0.047)	(0.073)	(0.067)	(0.059)	(0.102)	(0.092)	(0.073)	(0.300)	(0.115)
government job	0.146	0.029	0.354*	-0.025	0.046		-0.027	0.029	
creation program*	(0.118)	(0.188)	(0.142)	(0.306)	(0.324)		(0.314)	(0.302)	
spouse	0.087	-0.038	0.267*	0.218**	0.146	0.350**	0.233*	0.319	0.339
public sector*	(0.078)	(0.110)	(0.110)	(0.080)	(0.144)	(0.124)	(0.106)	(0.369)	(0.225)
spouse formal	0.016	0.104	0.024	0.002	0.059	0.022	-0.007	0.118	0.012
sector*	(0.056)	(0.083)	(0.081)	(0.073)	(0.111)	(0.117)	(0.081)	(0.277)	(0.116)
household	0.244***	0.262***	0.252**	0.084	0.134	0.060	0.119	0.323	0.149
public sector*	(0.048)	(0.066)	(0.078)	(0.089)	(0.126)	(0.178)	(0.108)	(0.323)	(0.221)
household	0.036	0.102	-0.034	0.079	0.200*	-0.044	0.097	0.453	-0.009
formal sector*	(0.040)	(0.059)	(0.056)	(0.062)	(0.092)	(0.114)	(0.074)	(0.242)	(0.109)
Pseudo R^2	0.252	0.203	0.329	0.292	0.210	0.399			
N	1438	675	753	871	444	423	830	435	394

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parenthesis. Language dummies are also included in the ordinary probit estimations. * after the name of the variable means that dy/dx is for discrete change of dummy variable from 0 to 1.

Table A.18 *Details of the Probit and IV Probit Estimation, Public vs. Informal Private, Asians*

	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
yschool	0.119***	0.092**	0.503*	0.138***	0.122***	0.143	0.125*	0.103	-0.001
	(0.027)	(0.028)	(0.202)	(0.031)	(0.036)	(2007.7)	(0.058)	(0.067)	(22.55)
female*	0.031			-0.046			-0.045		
	(0.115)			(0.157)			(0.175)		
age	0.041	0.012	-0.008	0.041	-0.014	0.092	0.028	-0.053	0.000
	(0.041)	(0.050)	(0.129)	(0.060)	(0.078)	(1316.6)	(0.065)	(0.088)	(9.750)
age^2/100	-0.047	-0.014	0.048	-0.045	0.016	-0.134	-0.033	0.057	0.000
	(0.051)	(0.060)	(0.175)	(0.071)	(0.090)	(1923.6)	(0.076)	(0.100)	(14.851)
married*	0.075	0.113	-0.105						
	(0.155)	(0.200)	(0.397)						
inverse dependecy	-0.116	-0.036	-0.493	0.446	0.402	-0.288	0.475	0.446	-0.007
ratio	(0.192)	(0.225)	(0.600)	(0.318)	(0.388)	(4072.5)	(0.393)	(0.447)	(206.1)
training*	0.090	0.215	-0.143	0.117	0.228	-0.998	0.102	0.234	0.000
	(0.180)	(0.232)	(0.423)	(0.179)	(0.206)	(32.776)	(0.195)	(0.189)	(0.160)
government job creation program*									
spouse	0.097	0.077		-0.619***	-0.566***		-0.618***	-0.609***	
public sector*	(0.344)	(0.390)		(0.087)	(0.090)		(0.130)	(0.141)	
spouse formal	-0.142	0.071	-0.555**	-0.324	0.012	-0.998	-0.345	0.002	0.000
sector*	(0.188)	(0.305)	(0.197)	(0.428)	(0.546)	(32.776)	(0.460)	(0.643)	(0.160)
household	0.284	0.245		0.381***	0.434***		0.382**	0.391**	
public sector*	(0.215)	(0.265)		(0.087)	(0.090)		(0.130)	(0.141)	
household	0.368**	0.266	0.357	0.334*	0.244	0.002	0.334*	0.252	1.000
formal sector*	(0.137)	(0.216)	(0.272)	(0.131)	(0.354)	(32.590)	(0.153)	(0.335)	(0.680)
Pseudo R^2	0.301	0.236	0.588	0.338	0.280	1.000			
N	151	101	40	104	74	20	98	71	19

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parenthesis. Language dummies are also included in the ordinary probit estimations. * after the name of the variable means that dy/dx is for discrete change of dummy variable from 0 to 1.

Table A.19 *Details of the Probit and IV Probit Estimation, Formal vs. Informal Private, All Races*

	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
yschool	0.013***	0.015***	0.021***	0.007***	0.007***	0.029***	0.015***	0.009***	0.056***
	(0.001)	(0.002)	(0.003)	(0.001)	(0.002)	(0.004)	(0.002)	(0.002)	(0.006)
female*	-0.152***			-0.208***			-0.256***		
	(0.011)			(0.020)			(0.017)		
age	0.006***	0.017***	0.001	0.002	0.008**	-0.004	0.001	0.005	-0.005
	(0.001)	(0.002)	(0.002)	(0.001)	(0.003)	(0.005)	(0.002)	(0.003)	(0.007)
age^2/100	-0.009***	-0.022***	-0.004	-0.003*	-0.011***	0.002	-0.003	-0.008*	0.006
	(0.001)	(0.003)	(0.002)	(0.001)	(0.003)	(0.006)	(0.002)	(0.004)	(0.008)
Black*	-0.174***	-0.155***	-0.308***	-0.137***	-0.107**	-0.351***	-0.229***	-0.205***	-0.363***
	(0.025)	(0.037)	(0.048)	(0.027)	(0.035)	(0.063)	(0.014)	(0.020)	(0.039)
Colored*	-0.089***	-0.074**	-0.178***	-0.087***	-0.086***	-0.197***	-0.082***	-0.080***	-0.128**
	(0.014)	(0.023)	(0.028)	(0.015)	(0.023)	(0.039)	(0.015)	(0.023)	(0.043)
Asian*	-0.052*	-0.055	-0.035	-0.050*	-0.056	-0.001	-0.023	-0.038	0.093
	(0.021)	(0.038)	(0.041)	(0.020)	(0.032)	(0.068)	(0.022)	(0.031)	(0.074)
married*	0.012**	0.073***	-0.087***						
	(0.004)	(0.011)	(0.014)						
inverse dependecy	0.046***	0.091***	0.043***	0.026***	0.049***	0.028	-0.005	-0.006	0.002
ratio	(0.007)	(0.013)	(0.011)	(0.006)	(0.012)	(0.022)	(0.010)	(0.016)	(0.032)
training*	0.028***	0.050***	0.026*	0.015**	0.026**	0.017	0.014	0.029*	-0.023
	(0.006)	(0.012)	(0.010)	(0.005)	(0.010)	(0.023)	(0.009)	(0.012)	(0.035)
government job	0.042***	-0.070	0.095***	0.006	-0.110	0.118**	0.003	-0.083	0.136
creation program*	(0.012)	(0.050)	(0.018)	(0.017)	(0.070)	(0.038)	(0.031)	(0.078)	(0.073)
spouse	0.010	-0.041	0.068***	0.029**	0.007	0.121***	0.051***	0.030	0.189**
public sector*	(0.013)	(0.035)	(0.016)	(0.010)	(0.030)	(0.037)	(0.015)	(0.032)	(0.058)
spouse formal	-0.024**	-0.033	0.029**	0.014*	0.012	0.060*	0.028**	0.023	0.127***
sector*	(0.009)	(0.021)	(0.011)	(0.006)	(0.014)	(0.025)	(0.010)	(0.017)	(0.036)
household	0.018*	0.034	0.022	-0.002	-0.005	0.007	-0.010	-0.012	-0.021
public sector*	(0.008)	(0.019)	(0.014)	(0.014)	(0.030)	(0.050)	(0.025)	(0.037)	(0.082)
household	0.073***	0.147***	0.076***	0.040***	0.070***	0.107***	0.065***	0.087***	0.120**
formal sector*	(0.008)	(0.017)	(0.013)	(0.007)	(0.014)	(0.024)	(0.009)	(0.013)	(0.037)
Pseudo R^2	0.188	0.132	0.237	0.217	0.121	0.281			
N	20428	11491	8937	10520	6466	4054	8983	5607	3376

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parenthesis. Language dummies are also included in the ordinary probit estimations. * after the name of the variable means that dy/dx is for discrete change of dummy variable from 0 to 1.

Table A.20 *Details of the Probit and IV Probit Estimation, Formal vs. Informal Private, Whites*

	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
yschool	0.016***	0.030***	0.004	0.010**	0.016**	0.003	0.014*	0.021*	0.010
	(0.005)	(0.008)	(0.003)	(0.004)	(0.006)	(0.007)	(0.006)	(0.009)	(0.014)
female*	0.001			-0.034			-0.045		
	(0.019)			(0.020)			(0.024)		
age	-0.006	-0.010	-0.003	-0.004	-0.007	-0.001	-0.006	-0.010	-0.004
	(0.006)	(0.010)	(0.005)	(0.005)	(0.008)	(0.010)	(0.006)	(0.009)	(0.013)
age^2/100	0.004	0.008	0.003	0.004	0.008	-0.001	0.006	0.011	0.003
	(0.007)	(0.012)	(0.005)	(0.006)	(0.008)	(0.012)	(0.007)	(0.010)	(0.015)
married*	0.042	0.123**	-0.050						
	(0.024)	(0.040)	(0.035)						
inverse dependecy	-0.004	0.030	-0.014	-0.012	-0.006	-0.027	-0.022	-0.019	-0.039
ratio	(0.026)	(0.049)	(0.018)	(0.021)	(0.031)	(0.041)	(0.026)	(0.038)	(0.052)
training*	-0.028	-0.006	-0.035	-0.011	0.022	-0.083	-0.013	0.027	-0.104
	(0.024)	(0.041)	(0.025)	(0.019)	(0.020)	(0.054)	(0.022)	(0.024)	(0.062)
government job creation program*									
spouse	-0.114	-0.808***	0.023	0.026	-0.933***	0.077	0.042	-0.915***	0.107*
public sector*	(0.134)	(0.046)	(0.027)	(0.043)	(0.022)	(0.045)	(0.044)	(0.018)	(0.052)
spouse formal	0.010	0.023	0.024*	0.025	0.034	0.008	0.041	0.047	0.038
sector*	(0.031)	(0.057)	(0.012)	(0.020)	(0.023)	(0.078)	(0.022)	(0.025)	(0.087)
household	0.087**	0.192***	0.015	0.005	0.066**	-0.042	0.003	0.084***	-0.061
public sector*	(0.032)	(0.046)	(0.029)	(0.056)	(0.023)	(0.135)	(0.070)	(0.018)	(0.164)
household	0.054*	0.106**	-0.002	0.023	0.023	0.053	0.027	0.030	0.059
formal sector*	(0.023)	(0.039)	(0.022)	(0.020)	(0.024)	(0.052)	(0.026)	(0.030)	(0.077)
Pseudo R^2	0.065	0.121	0.042	0.052	0.092	0.048			
N	2054	1209	845	1545	928	617	1517	912	605

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parenthesis. Language dummies are also included in the ordinary probit estimations. * after the name of the variable means that dy/dx is for discrete change of dummy variable from 0 to 1.

Table A.21 *Details of the Probit and IV Probit Estimation, Formal vs. Informal Private, Blacks*

	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
yschool	0.030***	0.023***	0.038***	0.022***	0.016***	0.035***	0.030***	0.015***	0.049***
	(0.001)	(0.002)	(0.003)	(0.002)	(0.003)	(0.006)	(0.003)	(0.004)	(0.005)
female*	-0.258***			-0.402***			-0.403***		
	(0.012)			(0.014)			(0.015)		
age	0.021***	0.032***	0.007	0.013**	0.027***	-0.001	0.009	0.019**	-0.003
	(0.003)	(0.004)	(0.004)	(0.004)	(0.006)	(0.007)	(0.005)	(0.007)	(0.007)
age^2/100	-0.029***	-0.042***	-0.014**	-0.019***	-0.035***	-0.001	-0.014*	-0.027***	0.004
	(0.003)	(0.004)	(0.005)	(0.005)	(0.007)	(0.009)	(0.006)	(0.008)	(0.009)
married*	0.026*	0.114***	-0.125***						
	(0.011)	(0.015)	(0.020)						
inverse dependecy	0.136***	0.153***	0.125***	0.127***	0.137***	0.098**	0.028	-0.001	0.053
ratio	(0.013)	(0.017)	(0.022)	(0.021)	(0.027)	(0.036)	(0.026)	(0.035)	(0.036)
training*	0.085***	0.082***	0.087***	0.049*	0.046	0.065	0.026	0.035	0.019
	(0.015)	(0.020)	(0.025)	(0.020)	(0.026)	(0.039)	(0.024)	(0.029)	(0.041)
government job	0.130***	-0.091	0.297***	0.017	-0.244*	0.211*	-0.002	-0.197	0.161
creation program*	(0.038)	(0.061)	(0.048)	(0.059)	(0.096)	(0.086)	(0.072)	(0.115)	(0.107)
spouse	0.013	-0.048	0.148**	0.100	0.020	0.195	0.146**	0.077	0.308*
public sector*	(0.037)	(0.054)	(0.053)	(0.053)	(0.087)	(0.100)	(0.053)	(0.082)	(0.136)
spouse formal	-0.074***	-0.054	0.043	0.037	0.022	0.075	0.059*	0.049	0.142*
sector*	(0.021)	(0.034)	(0.029)	(0.027)	(0.042)	(0.047)	(0.030)	(0.043)	(0.059)
household	0.019	0.039	0.007	-0.021	-0.005	-0.034	-0.063	-0.023	-0.111
public sector*	(0.025)	(0.036)	(0.035)	(0.058)	(0.081)	(0.078)	(0.071)	(0.085)	(0.079)
household	0.239***	0.295***	0.218***	0.170***	0.198***	0.185***	0.179***	0.206***	0.126*
formal sector*	(0.018)	(0.021)	(0.019)	(0.027)	(0.035)	(0.048)	(0.020)	(0.023)	(0.057)
Pseudo R^2	0.147	0.104	0.154	0.170	0.081	0.160			
N	14291	7971	6320	6647	4114	2533	5252	3323	1929

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parenthesis. Language dummies are also included in the ordinary probit estimations. * after the name of the variable means that dy/dx is for discrete change of dummy variable from 0 to 1.

Table A.22 *Details of the Probit and IV Probit Estimation, Formal vs. Informal Private, Coloreds*

	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
yschool	0.012***	0.001	0.036***	0.009**	0.001	0.038***	0.021***	0.000	0.072***
	(0.003)	(0.003)	(0.006)	(0.003)	(0.003)	(0.007)	(0.005)	(0.025)	(0.011)
female*	-0.148***			-0.240***			-0.285***		
	(0.021)			(0.039)			(0.030)		
age	-0.007*	-0.004	-0.015	-0.009	-0.004	-0.021	-0.011	-0.028	-0.014
	(0.004)	(0.005)	(0.008)	(0.006)	(0.008)	(0.016)	(0.007)	(0.038)	(0.018)
age^2/100	0.005	0.001	0.009	0.008	0.002	0.017	0.010	0.018	0.014
	(0.004)	(0.006)	(0.010)	(0.006)	(0.009)	(0.019)	(0.009)	(0.043)	(0.021)
married*	0.005	0.048*	-0.120**						
	(0.015)	(0.023)	(0.040)						
inverse dependecy	-0.039*	-0.012	-0.074	-0.051	-0.033	-0.126	-0.068	-0.204	-0.074
ratio	(0.020)	(0.029)	(0.044)	(0.026)	(0.037)	(0.066)	(0.036)	(0.185)	(0.077)
training*	0.031	0.061*	-0.003	0.060*	0.084*	0.055	0.071*	0.581*	-0.014
	(0.020)	(0.030)	(0.050)	(0.024)	(0.036)	(0.081)	(0.033)	(0.239)	(0.099)
government job	-0.143	-0.179	-0.070	-0.002	-0.127	-0.127	-0.006	-0.083	
creation program*	(0.114)	(0.169)	(0.186)	(0.123)	(0.292)		(0.168)	(0.297)	
spouse	0.022	-0.040	0.141**	0.064*	0.033	0.199*	0.094*	0.240	0.197
public sector*	(0.033)	(0.068)	(0.054)	(0.031)	(0.057)	(0.095)	(0.044)	(0.339)	(0.146)
spouse formal	-0.003	0.000	0.096**	0.020	0.004	0.141*	0.027	0.042	0.186*
sector*	(0.020)	(0.033)	(0.033)	(0.022)	(0.036)	(0.060)	(0.032)	(0.181)	(0.075)
household	0.043*	0.016	0.118**	0.002	-0.070	0.127	-0.007	-0.294	0.158
public sector*	(0.020)	(0.035)	(0.043)	(0.041)	(0.080)	(0.098)	(0.062)	(0.286)	(0.139)
household	0.080***	0.105***	0.109***	0.073***	0.090**	0.103	0.111***	0.676***	0.117
formal sector*	(0.017)	(0.030)	(0.031)	(0.022)	(0.034)	(0.065)	(0.022)	(0.158)	(0.085)
Pseudo R^2	0.113	0.073	0.186	0.135	0.083	0.169			
N	3473	1895	1569	1895	1116	762	1815	1098	717

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parenthesis. Language dummies are also included in the ordinary probit estimations. * after the name of the variable means that dy/dx is for discrete change of dummy variable from 0 to 1.

Table A.23 *Details of the Probit and IV Probit Estimation, Formal vs. Informal Private, Asians*

	Probit			Probit (married people)			IV Probit		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
yschool	0.039**	0.038**	0.037	0.041***	0.041**	0.064	0.026	0.030	0.031
	(0.012)	(0.014)	(0.021)	(0.012)	(0.013)	(0.038)	(0.022)	(0.022)	(0.081)
female*	0.025			0.006			0.011		
	(0.054)			(0.059)			(0.059)		
age	0.001	-0.002	0.015	-0.012	-0.033	0.068	-0.016	-0.041*	0.072
	(0.016)	(0.020)	(0.027)	(0.019)	(0.023)	(0.065)	(0.019)	(0.020)	(0.077)
age^2/100	0.003	0.007	-0.015	0.018	0.040	-0.064	0.020	0.047*	-0.075
	(0.019)	(0.024)	(0.034)	(0.023)	(0.027)	(0.077)	(0.022)	(0.024)	(0.094)
married*	0.044	0.058	-0.038						
	(0.064)	(0.081)	(0.131)						
inverse dependecy	-0.191*	-0.189	-0.219	-0.038	-0.051	-0.048	-0.057	-0.078	-0.050
ratio	(0.081)	(0.097)	(0.150)	(0.091)	(0.106)	(0.282)	(0.103)	(0.115)	(0.341)
training*	0.060	0.093	-0.046	0.017	0.050	-0.128	0.021	0.050	-0.128
	(0.070)	(0.078)	(0.154)	(0.067)	(0.067)	(0.242)	(0.069)	(0.061)	(0.255)
government job creation program*									
spouse	0.005	-0.112		-0.850***	-0.855***		-0.854***	-0.878***	
public sector*	(0.198)	(0.281)		(0.091)	(0.034)		(0.039)	(0.039)	
spouse formal	-0.042	-0.037	-0.013	-0.134	-0.087	-0.766	-0.112	-0.072	-0.700**
sector*	(0.108)	(0.133)	(0.175)	(0.189)	(0.177)	(0.691)	(0.188)	(0.166)	(0.219)
household	0.153*		0.138	0.150***	0.145***		0.146***	0.122**	
public sector*	(0.075)	(0.098)		(0.034)	(0.034)		(0.039)	(0.039)	
household	0.177***	0.188**	0.152	0.139***	0.131***	0.233	0.134***	0.112**	0.300
formal sector*	(0.053)	(0.066)	(0.079)	(0.034)	(0.036)	(0.143)	(0.040)	(0.038)	(0.219)
Pseudo R^2	0.175	0.166	0.199	0.175	0.168	0.286			
N	578	389	165	408	285	102	391	270	102

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors are in the parenthesis. Language dummies are also included in the ordinary probit estimations. * after the name of the variable means that dy/dx is for discrete change of dummy variable from 0 to 1.